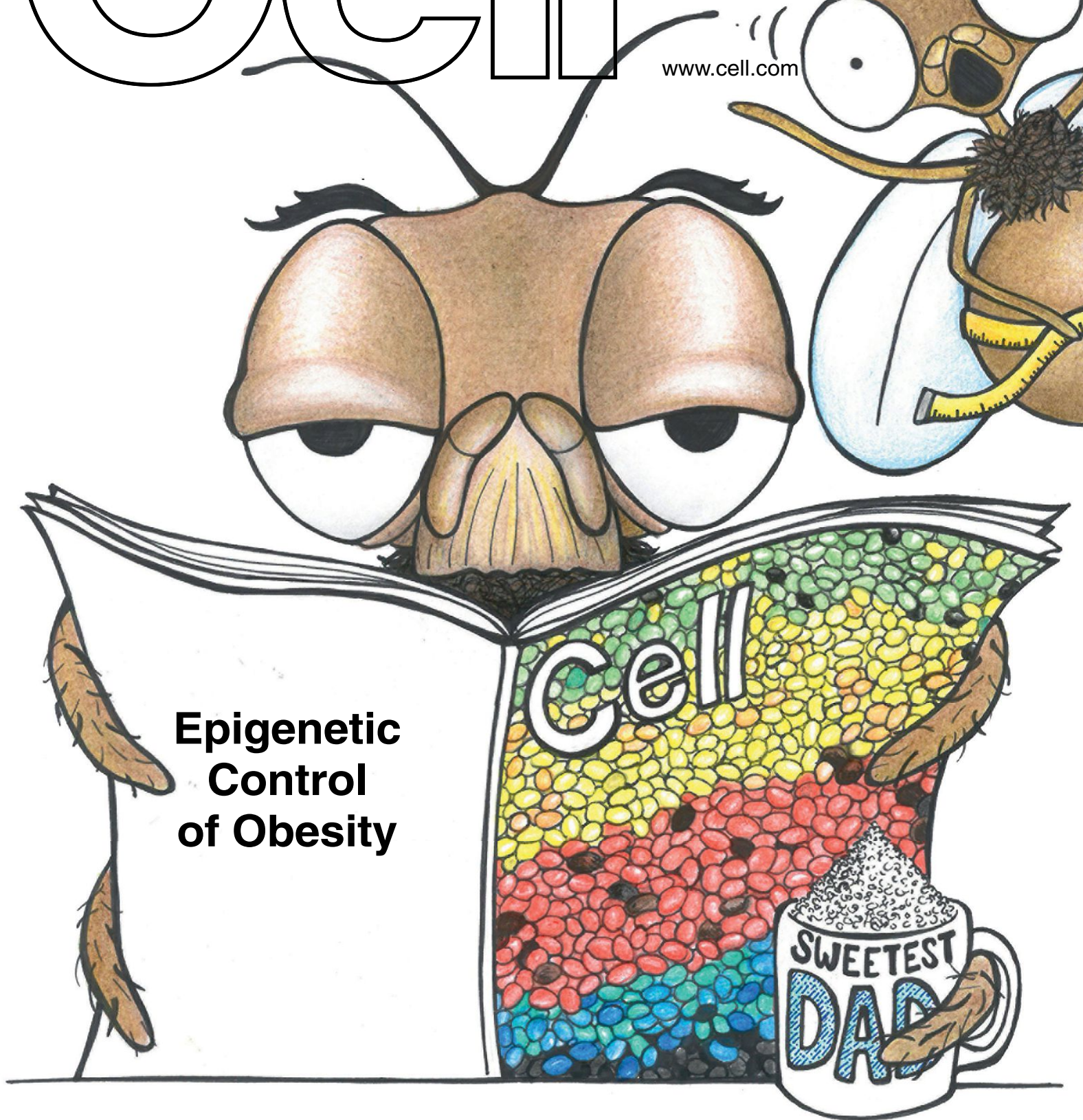


Cell

Volume 159
Number 6

December 4, 2014

www.cell.com



**Epigenetic
Control
of Obesity**

more power to you.

Thermo Scientific™ Pierce™ Power System lets you power through two essential protein detection steps with one efficient system! Save hours in coomassie staining and protein transfer from gels to membranes without compromising the quality of your results. Efficiently stain or transfer protein gels in about 10 minutes with one system. Intuitive, color LCD touchscreen interface and pre-programmed methods make both coomassie staining and protein transfer easy. One power station + two cassettes = powerful possibilities.

• Request a demo at thermoscientific.com/powersystem



Thermo Scientific Pierce Power Stainer

Coomassie staining and destaining of proteins in about 10 minutes



Thermo Scientific Pierce Power Blotter

Transfer proteins in 5-10 minutes





Smooth Operator

The new Eppendorf micromanipulators **TransferMan® 4r** and **InjectMan® 4**

The new Eppendorf micromanipulators have been developed to simplify your work and increase the quality of your micromanipulation results. Based on their ergonomic design, the robust and reliable devices provide excellent performance for a wide range of applications.

- > Unprecedented movement control and maximum stability
- > Pre-defined functions for various applications to optimize your workflow
- > Connection with new Eppendorf electronic microinjectors **FemtoJet® 4i/4x** and **Eppendorf PiezoXpert®**

www.eppendorf.com/cellmanipulation

F · S · T[®]
FINE SCIENCE TOOLS

Precision in Its Own Words

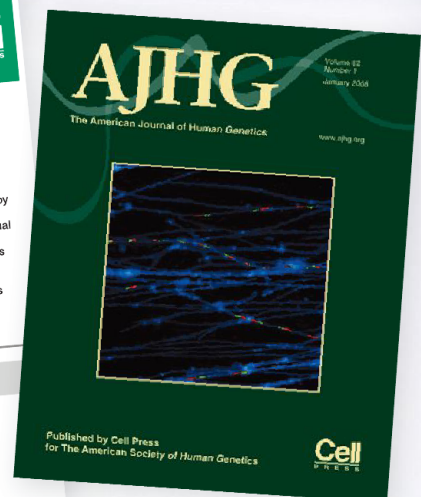
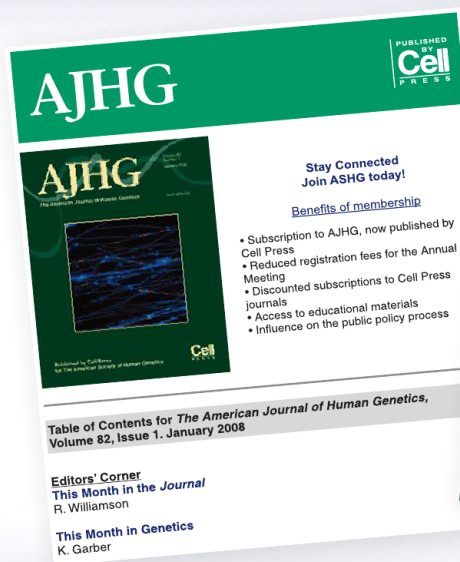
The extraordinary quality of our surgical and microsurgical instruments is the result of our relentless attention to detail. Every instrument we sell is designed to exacting specifications, forged from the strongest, lightest materials available, and tested to ensure precision performance.

FINE SURGICAL INSTRUMENTS FOR RESEARCH™
Visit us at finescience.com or call **800 521 2109**



Be the first...

to read the latest issue of AJHG.



Receive AJHG Email Alerts – FREE!
Register Now at www.ajhg.org

Cell

Editor

Emilie Marcus

Senior Deputy Editor

Elena Porro

Deputy Editors

Karen Carniol
Robert Kruger

Scientific Editors

Mirna Kvafo
Cindy C. Lu
Steve Mao
João Monteiro
Sri Devi Narasimhan
Lara Szewczak
Jiaying Tan

Senior Managing Editor

Andy Smith

Managing Editor

Trina Arpin

Deputy Managing Editor

Julie Fiorilla

Art Program Manager

Andrew A. Tang

Managing Designer and Illustrator

Yvonne Blanco

Illustrator

Kip B. Lyall

Editorial Assistant

Jared Graves

Editorial Board

C. David Allis
Geneviève Almouzni
Angelika Amon
Silvia Arber
Johan Auwerx
Richard Axel
Cori Bargmann
Konrad Basler
David Baulcombe
Yasmine Belkaid
Jeffrey Benovic
Wendy Bickmore
Elizabeth Blackburn
Joan Brugge
Lewis Cantley
Xuetao Cao
Joanne Chory
David Clapham
Andrew Clark
Hans Clevers
Pascale Cossart
Benjamin Cravatt
George Daley
Jeff Dangl
Ted Dawson
Karl Deisseroth
Hongkui Deng
Pier Paolo di Fiore
Ivan Dikic
Julian Downward
Bruce Edgar
Steve Elledge
Anne Ephrussi
Ronald Evans
Witold Filipowicz
Marco Foiani
Elaine Fuchs
Daniel Geschwind

Yukiko Goda
Stephen Goff
Joe Goldstein
Douglas Green
Leonard Guarente
Taekjip Ha
Daniel Haber
Ulrike Heberlein
Mark Hochstrasser
Erika Holzbaur
Huck Hui Ng
Tony Hunter
James Hurley
Richard Hynes
Trey Ideker
Thomas Jessell
Tarun Kapoor
Narry Kim
Mary-Claire King
David Kingsley
Frank Kirchhoff
Robert Lamb
Mark Lemmon
Beth Levine
Wendell Lim
Jennifer Lippincott-Schwartz
Dan Littman
Richard Losick
Scott Lowe
Tom Maniatis
Matthias Mann
Kelsey Martin
Joan Massagué
Satyajit Mayor
Ruslan Medzhitov
Craig Mello
Tom Misteli
Tim Mitchison
Danesh Moazed

Alex Mogilner
Paul Nurse
Roy Parker
Dana Pe'er
Kathrin Plath
Carol Prives
Klaus Rajewsky
Lalita Ramakrishnan
Venki Ramakrishnan
Rama Ranganathan
Anne Ridley
Alexander Rudensky
Helen Saibil
Charles Sawyers
Ueli Schibler
Joseph Schlessinger
Hans Schöler
Trina Schroer
Kevan Shokat
Nahum Sonenberg
James Spudich
Deepak Srivastava
Azim Surani
Keiji Tanaka
Craig Thompson
Robert Tjian
Ulrich von Andrian
Gerhard Wagner
Jonathan Weissman
Matthew Welch
Tian Xu
Shinya Yamanaka
Richard Young
Marino Zerial
Xiaowei Zhuang
Juleen Zierath
Huda Zoghbi

Cell Office

Cell, Cell Press, 600 Technology Square, 5th Floor, Cambridge, Massachusetts 02139
Phone: (+1) 617 661 7057, Fax: (+1) 617 661 7061, E-mail: celleditor@cell.com
Online Publication: <http://www.cell.com>

Cell (ISSN 0092-8674) is published biweekly by Cell Press, 600 Technology Square, 5th Floor, Cambridge, Massachusetts 02139. The institutional subscription rate for 2014 is \$1,682 (US and Canada) or \$1,895 (elsewhere). The individual subscription rate is \$246 (US and Canada) or \$371 (elsewhere). The individual copy price is \$50. Periodicals postage paid at Boston, Massachusetts and additional mailing offices. Postmaster: send address changes to Elsevier Customer Service Americas, Cell Press Journals, 3251 Riverport Lane, Maryland Heights, MO 63043, USA.

ⒸThe paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper). Printed by Dartmouth Printing Company, Hanover, NH.

A Chance to Expand Your Creative Possibilities



IBS invites you to join the challenge of boosting basic science.

IBS (Institute for Basic Science) is a government-funded research institute, established with the sole purpose of driving forward the development of basic science in Korea. We have launched 21 research centers; each headed by an internationally renowned scientist with an operational budget of up to 10 million USD a year. IBS will be comprised of a total of 50 research centers in all fields of basic science, including mathematics, physics, chemistry, life science, earth science and interdisciplinary science.

Our strategy is simple: establish an excellent research institute for basic science in Korea by recruiting exceptional scientists worldwide and fully supporting them in whichever research areas they wish to pursue for the long-term. IBS ensures that we select the best possible people for this challenge with our peer-review evaluation system. You are cordially invited to join IBS to boost basic science to the next level.

Recruitment for IBS

IBS looks forward to shaping the future of basic science with you. Visit our website for detailed recruitment information and register yourself with the IBS Talent Pool DB for a future career with IBS.

<http://www.ibs.re.kr/en> <http://www.ibs.re.kr/talents>

Cell Press

CEO

Emilie Marcus

Vice President of Business Development

Joanne Tracy

Vice President of Operations

Keith Wollman

Publishing Directors

Peter Lee

Deborah Sweet

Editorial Director, Reviews Strategy

Katja Brose

Editorial Director, Content Development

Elena Porro

Director of Production

Meredith Adinolfi

Director of Marketing

Jonathan Atkinson

Media Relations Manager

Mary Beth O'Leary

Display Advertising

For advertising, sponsorship, and reprint opportunities, contact Jonathan Christison (jchristison@cell.com) or visit <http://advertisers.cell.com/contact-us>.

Recruitment Advertising

For recruitment advertising, visit <http://recruitmentads.cell.com/contact-us>.

©2014 Elsevier Inc. All rights reserved.

This journal and the individual contributions contained in it are protected under copyright by Elsevier Inc., and the following terms and conditions apply to their use:

Photocopying

Single photocopies of single articles may be made for personal use as allowed by national copyright laws. Permission of the Publisher and payment of a fee is required for all other photocopying, including multiple or systematic copying, copying for advertising or promotional purposes, resale, and all forms of document delivery. Special rates are available for educational institutions that wish to make photocopies for nonprofit educational classroom use.

Permissions

For information on how to seek permission, visit <http://www.elsevier.com/permissions> or call +44 1865 843830 (UK) / +1 215 239 3804 (US).

Derivative Works

Subscribers may reproduce tables of contents or prepare lists of articles including summaries for internal circulation within their institutions. Permission of the Publisher is required for resale or distribution outside the institution. Permission of the Publisher is required for all other derivative works, including compilations and translations (please consult <http://www.elsevier.com/permissions>).

Electronic Storage or Usage

Permission of the Publisher is required to store or use electronically any material contained in this journal, including any article or part of an article (please consult www.elsevier.com/permissions). Except as outlined above, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without prior written permission of the Publisher.

Notice

No responsibility is assumed by the Publisher for any injury and/or damage to persons or property as a matter of products liability, negligence, or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made. Although all advertising material is expected to conform to ethical (medical) standards, inclusion in this publication does not constitute a guarantee or endorsement of the quality or value of such product or of the claims made of it by its manufacturer.

Reprints

Article reprints are available through Elsevier's reprint service; for information, contact Derrick Imasa (e-mail: d.imasa@elsevier.com; ph: +1 212 633 3874).

Subscription Orders and Inquiries

Mail, fax, or e-mail address changes to Elsevier Customer Service Americas, allowing 4–6 weeks for processing. Lost or damaged issues will be replaced, subject to availability, if Cell Press is notified within the claim period (US and airmail delivery: 3 months from issue date; surface delivery: 4 months from issue date). Periodical delivery in the US can take up to 3 weeks. Airmail delivery can take 2–4 weeks. Mailing address: Elsevier Customer Service Americas, Cell Press Journals, 3251 Riverport Lane, Maryland Heights, MO 63043, USA. Toll-free phone within USA/Canada: +1 866 314 2355; phone for outside US/Canada: +1 314 447 8880; fax: +1 314 447 8029; e-mail: subs@cell.com.

Funding Body Agreements and Policies

Elsevier has established agreements and developed policies to allow authors whose articles appear in journals published by Elsevier to comply with potential manuscript archiving requirements as specified as conditions of their grant awards. To learn more about existing agreements and policies, please visit <http://www.cell.com/cellpress/FundingBodyAgreements>.

NEW

transEDIT™ Lentiviral CRISPR-Cas9 Reagents

*Optimized gRNA designs,
versatile vectors, flexible formats
for efficient gene editing*



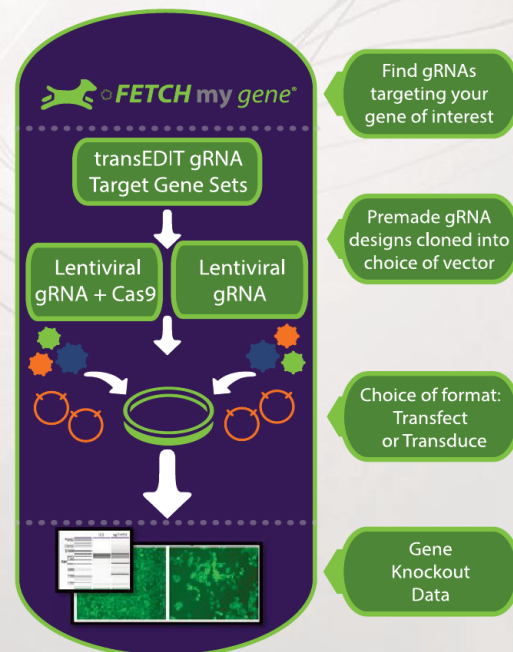
Perform efficient genome editing with transEDIT™ lentiviral CRISPR-Cas9 reagents. Simply search for your gene or target sequence of interest to order top ranked gRNA designs in your choice of lentiviral gRNA and Cas9 vectors (including custom inducible Cas9 options). transEDIT lentiviral vectors enable optimal expression and selection of gRNA and Cas9 for high efficiency DNA cleavage and successful gene editing.

- All-in-one (gRNA plus Cas9) or single gRNA delivery with inducible options
- Single or paired guide RNA strategies for gene editing
- Multiple markers enable selection for high expressing cells
- Pooled gRNA libraries targeting gene families and pathways

transEDIT CRISPR-Cas9 reagents enable specific gene knockouts, targeted insertions and multiplexed screening of gene families or pathways.

New! Pooled gRNA library
targeting chromatin regulators

SIMPLE - FLEXIBLE - FAST



For more information and to order see www.transomic.com/transEDIT-CRISPR



Leading Edge

Cell Volume 159 Number 6, December 4, 2014

IN THIS ISSUE

SELECT

1235 Biological Code Breaking

BENCHMARKS

1239 A Place and a Grid in the Sun

E. Kandel

1243 Better Imaging through Chemistry

E.H.K. Stelzer

PREVIEWS

1247 Speciation via Autoimmunity:
A Dangerous Mix

N. Phadnis and H.S. Malik

1249 A Split Personality for Nucleosomes

D.J. McKay and J.D. Lieb

1251 How TriC Folds Tricky Proteins

A. Zhuravleva and S.E. Radford

PERSPECTIVE

1253 Metabolic Inflexibility: When Mitochondrial
Indecision Leads to Metabolic Gridlock

D.M. Muoio

REVIEW

1263 Metabolic Control of Autophagy

L. Galluzzi, F. Pietrocola, B. Levine, and G. Kroemer

SNAPSHOT

1478 Hormones of the Gastrointestinal Tract

K.C. Coate, S.A. Kliewer, and D.J. Mangelsdorf

(continued)

A smarter approach to FLUORESCENCE IMAGING



Introducing the ZOE™ Fluorescent Cell Imager.
No darkroom, no training, no overwhelming user interface.

Combining brightfield capabilities with multichannel fluorescence, this cell imager is both affordable and easy to use — your perfect solution for routine cell culture and imaging applications.

Learn more at [bio-rad.com/info/newzoe](https://www.bio-rad.com/info/newzoe)

BIO-RAD

Articles

Cell Volume 159 Number 6, December 4, 2014

- 1277** Gut Microbiota Elicits a Protective Immune Response against Malaria Transmission
B. Yilmaz, S. Portugal, T.M. Tran, R. Gozzelino, S. Ramos, J. Gomes, A. Regalado, P.J. Cowan, A.J.F. d'Apice, A.S. Chong, O.K. Doumbo, B. Traore, P.D. Crompton, H. Silveira, and M.P. Soares
- 1290** Host Adaptation of a Bacterial Toxin from the Human Pathogen *Salmonella* Typhi
L. Deng, J. Song, X. Gao, J. Wang, H. Yu, X. Chen, N. Varki, Y. Naito-Matsui, J.E. Galán, and A. Varki
- 1300** Beta-Lactam Antibiotics Induce a Lethal Malfunctioning of the Bacterial Cell Wall Synthesis Machinery
H. Cho, T. Uehara, and T.G. Bernhardt
- 1312** Tissue-Resident Macrophage Enhancer Landscapes Are Shaped by the Local Microenvironment
Y. Lavin, D. Winter, R. Blecher-Gonen, E. David, H. Keren-Shaul, M. Merad, S. Jung, and I. Amit
- 1327** Environment Drives Selection and Function of Enhancers Controlling Tissue-Specific Macrophage Identities
D. Gosselin, V.M. Link, C.E. Romanoski, G.J. Fonseca, D.Z. Eichenfield, N.J. Spann, J.D. Stender, H.B. Chun, H. Garner, F. Geissmann, and C.K. Glass
- 1341** Species-wide Genetic Incompatibility Analysis Identifies Immune Genes as Hot Spots of Deleterious Epistasis
E. Chae, K. Bomblies, S.-T. Kim, D. Karelina, M. Zaidem, S. Ossowski, C. Martín-Pizarro, R.A.E. Laitinen, B.A. Rowan, H. Tenenboim, S. Lechner, M. Demar, A. Habring-Müller, C. Lanz, G. Rättsch, and D. Weigel
- 1352** Paternal Diet Defines Offspring Chromatin State and Intergenerational Obesity
A. Öst, A. Lempradl, E. Casas, M. Weigert, T. Tiko, M. Deniz, L. Pantano, U. Boenisch, P.M. Itskov, M. Stoeckius, M. Ruf, N. Rajewsky, G. Reuter, N. Iovino, C. Ribeiro, M. Alenius, S. Heyne, T. Vavouri, and J.A. Pospisilik
- 1365** Uridylation by TUT4 and TUT7 Marks mRNA for Degradation
J. Lim, M. Ha, H. Chang, S.C. Kwon, D.K. Simanshu, D.J. Patel, and V.N. Kim
- 1377** Subnucleosomal Structures and Nucleosome Asymmetry across a Genome
H.S. Rhee, A.R. Bataille, L. Zhang, and B.F. Pugh
- 1389** Proteostatic Control of Telomerase Function through TRiC-Mediated Folding of TCAB1
A. Freund, F.L. Zhong, A.S. Venteicher, Z. Meng, T.D. Veenstra, J. Frydman, and S.E. Artandi
- 1404** Leptin Mediates the Increase in Blood Pressure Associated with Obesity
S.E. Simonds, J.T. Pryor, E. Ravussin, F.L. Greenway, R. Dileone, A.M. Allen, J. Bassi, J.K. Elmquist, J.M. Keogh, E. Henning, M.G. Myers, Jr., J. Licinio, R.D. Brown, P.J. Enriori, S. O'Rahilly, S.M. Sternson, K.L. Grove, D.C. Spanswick, I.S. Farooqi, and M.A. Cowley

(continued)



ANTIBODY WISH LIST

- ✓ WORKS IN ALL MY APPLICATIONS
- ✓ HIGH SENSITIVITY AND BINDING AFFINITY
- ✓ EXTENSIVELY VALIDATED IN UNMODIFIED SAMPLES
- ✓ ALL VALIDATION DATA AVAILABLE ONLINE
- ✓ SUITABLE FOR USE IN MULTIPLE SPECIES
- ✓ SATISFACTION GUARANTEED OR MY GRANT MONEY BACK
- ✓ 100% ORIGINAL PRODUCT FROM THE ORIGINAL MANUFACTURER



PROTEINTech GROUP FULFILS EVERY
SCIENTIST'S ANTIBODY WISH LIST



Find out more at
ptglab.com

1417 Identification of Spinal Circuits
Transmitting and Gating Mechanical Pain

*B. Duan, L. Cheng, S. Bourane, O. Britz, C. Padilla,
L. Garcia-Campmany, M. Krashes, W. Knowlton,
T. Velasquez, X. Ren, S.E. Ross, B.B. Lowell, Y. Wang,
M. Goulding, and Q. Ma*

1433 A Constant Size Extension Drives
Bacterial Cell Size Homeostasis

*M. Campos, I.V. Surovtsev, S. Kato, A. Paintdakhi,
B. Beltran, S.E. Ebmeier, and C. Jacobs-Wagner*

1447 The Structure and Regulation
of Human Muscle α -Actinin

*E.d.A. Ribeiro, Jr., N. Pinotsis, A. Ghisleni, A. Salmazo,
P.V. Konarev, J. Kostan, B. Sjöblom, C. Schreiner,
A.A. Polyansky, E.A. Gkougkoulia, M.R. Holt,
F.L. Aachmann, B. Žagrović, E. Bordinon, K.F. Pirker,
D.I. Svergun, M. Gautel, and K. Djinović-Carugo*

RESOURCE

1461 Integration of Genomic Data Enables
Selective Discovery of Breast Cancer Drivers

*F. Sanchez-Garcia, P. Villagrasa, J. Matsui, D. Kotliar,
V. Castro, U.-D. Akavia, B.-J. Chen, L. Saucedo-Cuevas,
R. Rodriguez Barrueco, D. Llobet-Navas, J.M. Silva,
and D. Pe'er*

ERRATUM

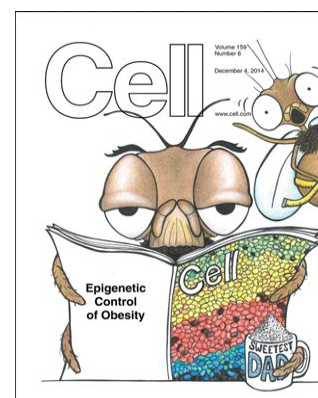
1476 TRIP12 and UBR5 Suppress
Spreading of Chromatin Ubiquitylation
at Damaged Chromosomes

*T. Gudjonsson, M. Altmeyer, V. Savic, L. Toledo, C. Dinant,
M. Grøfte, J. Bartkova, M. Poulsen, Y. Oka,
S. Bekker-Jensen, N. Mailand, B. Neumann, J.-K. Heriche,
R. Shearer, D. Saunders, J. Bartek, J. Lukas, and C. Lukas*

ANNOUNCEMENTS

POSITIONS AVAILABLE

On the cover: The global rise in obesity has warranted a need for identifying genetic and epigenetic factors underlying the disease. Öst et al. (pp. 1352–1364) present a *Drosophila* model of paternal-diet-induced metabolic reprogramming where altered sugar intake in fathers leads to obesity in the offspring. Paternal sugar acts to modify germline and zygotic chromatin states as well as the transcription of metabolic genes. This altered chromatin signature is observed to be conserved in mouse and humans and may help to explain inter-generational regulation of metabolic states. Image by Gabriela Navarrete Bovio.



FORWARD FROM

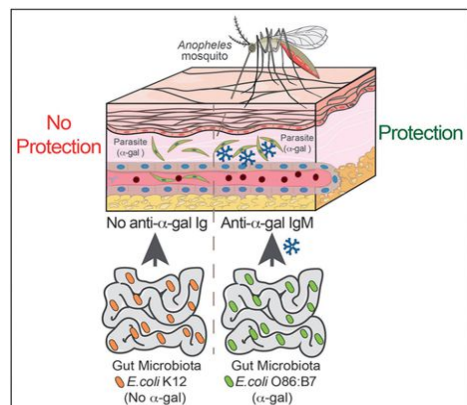
40



As we commemorate 40 years of publishing exciting biology we're also embracing the future. We will continue to lead the way with best-in-class editorial service, innovations in publishing, and new technology to address the evolving needs of scientists. Here's to the next 40 years!

www.cell.com/40/home

Cell



Microbiome versus Malaria

PAGE 1277

Yilmaz et al. find that the bacteria, found within the gut microbiome, carrying a specific glycan residue elicit antibodies in humans to an epitope also found on *Plasmodium* parasites. Presence of the antibodies correlates with protection from malaria infection in a human cohort and provides sterile protection following vaccination in a mouse model of infection.

Sugar Opens the Door for Typhoid

PAGE 1290

Salmonella Typhi causes typhoid fever exclusively in humans and depends on binding of the typhoid toxin to a specific glycan receptor on target cells. Deng et al. discover that the toxin preferentially binds glycans terminated in N-acetylneuraminic acid (Neu5Ac) and that humans lack the enzyme, which in other mammals converts Neu5Ac to N-glycolyneuraminic acid.

The Home Makes the Macrophage

PAGE 1312 and PAGE 1327

Macrophages are cells of the innate immune system that also contribute to the homeostasis of distinct tissues in the organism. Lavin et al. and Gosselin et al. now show that the tissue environment can alter the chromatin landscape of resident macrophages to equip them with functions that are specific and adequate to each tissue. Surprisingly, tissue-resident terminally-differentiated cells can be reprogrammed by transplantation into a new tissue, demonstrating a high degree of plasticity and highlighting the role of the environment in shaping macrophage function.

Plant Immunity Shapes Genome Evolution

PAGE 1341

Genetic incompatibilities prevent specific alleles from assembling into single genotypes, influencing genome and species-wide patterns of sequence variation. Chae et al. now show that incompatibility hotspots in *Arabidopsis* are enriched in immune receptor genes. When these genes interact with each other in hybrids, conflict within the immune system often arises, limiting the number of viable combinations of disease resistance alleles, thereby shaping the evolution of the plant genome.

mRNA Destr-U-ct Code

PAGE 1365

Lim and colleagues discover that terminal uridylation serves as a general molecular mark of mRNA decay. Two enzymes, TUT4 and TUT7, facilitate uridylation in mammalian cells by selectively recognizing mRNAs with shortened poly(A) tails and inducing decay of the deadenylated mRNA body.

Chaperone Cen-TRiC View for Telomeres

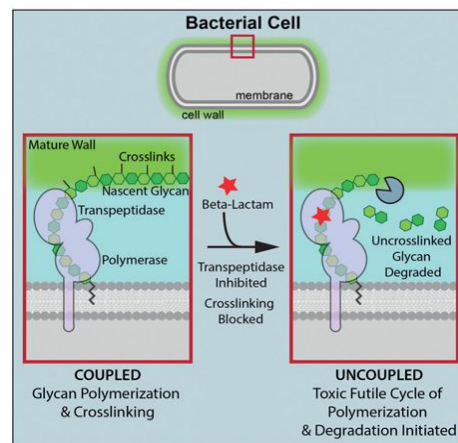
PAGE 1389

Telomerase function depends on a complex pathway for assembly, localization, and association with telomeres. Freund et al. identify the chaperonin CCT/TRiC as being required for telomerase trafficking and telomere elongation. Mutations in a telomerase cofactor that impair folding by TRiC cause dyskeratosis congenita, thus establishing a critical role for proteostatic control of telomere function.

Penicillin, We Thought We Knew Ya'

PAGE 1300

Penicillin is thought to inhibit bacterial cell wall synthesis, but Cho et al. now demonstrate that beta lactam antibiotics induce a toxic, futile cycle of peptidoglycan (PG) synthesis and degradation that depletes cellular stores of the PG building blocks. Additionally, characterization of the futile cycle revealed a quality control mechanism in cell wall biosynthesis.



Inspire your peers with Cell Reports



Breaking new ground? From molecular genetics to developmental neurobiology, *Cell Reports* publishes thought-provoking, cutting edge research spanning the spectrum of life sciences.

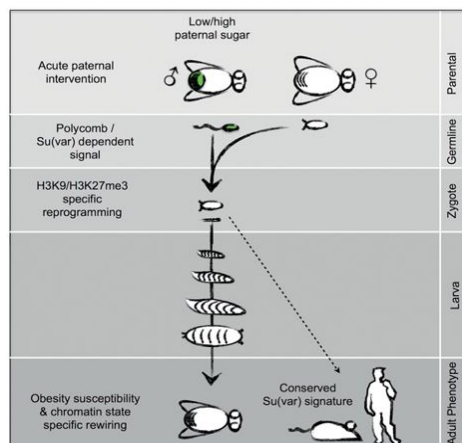
Cell Reports is an open access journal that offers the quality, rigor, and visibility you've come to expect from Cell Press.

Do you have a new biological insight? Submit your paper to *Cell Reports*.

For more information visit www.cell.com/cell-reports

An open access journal with impact.
From Cell Press.

Cell
Reports



No Sugar for Daddy

PAGE 1352

Öst and colleagues find that a brief high-sugar dietary intervention in *Drosophila* fathers controls offspring adiposity by reprogramming select chromatin regions in the offspring. Paternal sugar intake relieves silencing of select chromatin regions and leads to reprogramming of metabolic genes during two distinct germline and zygotic windows. Importantly, this altered gene signature is conserved in mice and humans and may explain mechanisms underlying inter-generational metabolic control.

Leptin Boosts Blood Pressure

PAGE 1404

Leptin, an adipose-derived hormone, is elevated in obesity. Simonds et al. find a role for leptin in the brain, modulating blood pressure by activating neurons in the dorsomedial nucleus of the hypothalamus. Blocking leptin action leads to

reduced blood pressure in obese mice, and clinical studies support this role for leptin in mediating the increased risk of high blood pressure associated with obesity.

The Pain Locker

PAGE 1417

Providing insight into the basis for sensing pain and developing new treatments, Duan et al. identify the spinal cord neurons that transmit and gate pain. They demonstrate that nociceptors and mechanoreceptors provide input to a population of somatostatin-positive excitatory neurons that are in turn gated by dynorphin-inhibitory neurons.

One Growth Spurt Fits All

PAGE 1433

Bacterial size homeostasis is thought to be achieved through a threshold mechanism that promotes division when cells reach a critical size. Campos et al. now show that bacteria instead elongate by a fixed amount between divisions, irrespective of their size at birth. This set elongation leads to robust cell size homeostasis and provides new governing rules for cell-cycle control.

From Actinin to Z in Muscle

PAGE 1447

Sarcomeres are the basic contractile units of striated muscles, and the Z disk region within them plays a key role in maintaining cellular architecture and signalling. Ribeiro et al. present a three dimensional structure of α -Actinin, the major component of the Z disk, that explains how a phosphoinositide-based regulation mechanism allows it bind to titin and control Z disk assembly.

Shedding Light on CNVs in Cancer

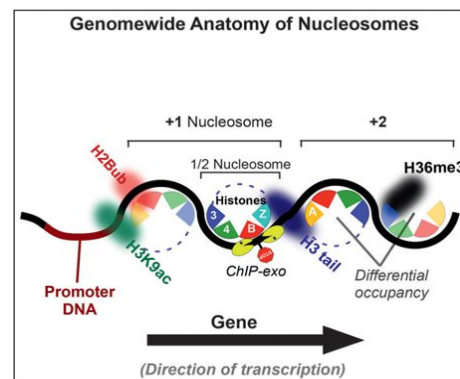
PAGE 1461

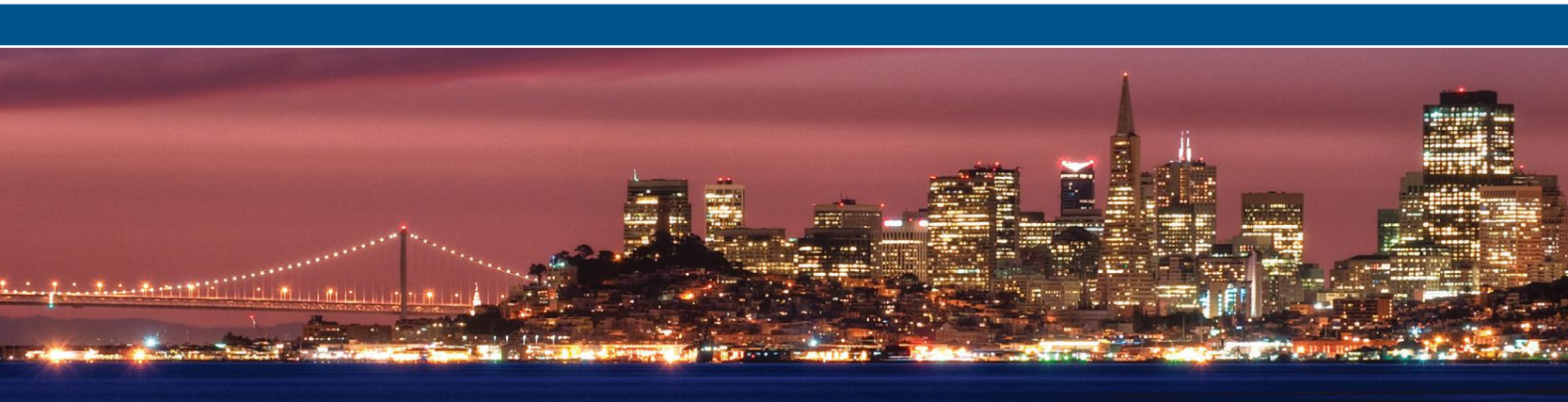
Discerning which copy-number variants drive cancer has been a challenging problem. By developing and validating a new computational algorithm called HELIOS, which integrates genomic data from primary tumors with outcomes of functional RNAi screens, Sanchez-Garcia et al. double the known landscape of breast cancer drivers.

Not Just Octamers on a String

PAGE 1377

Genes are packaged into nucleosomes that are generally characterized as octamers containing different histone subtypes and variants. Rhee et al. perform a ChIP-exo analysis of histone-DNA contacts near promoters and find unexpected nucleosome modularity, including evidence for existence of subnucleosomal particles such as tetramers as well as asymmetric patterns of histone modifications within a single nucleosome that correlate with the direction of transcription.





Abstract
submission
deadline:

August 22
2014

Early bird
deadline:

October 3
2014

Stem Cell Energetics will bring together academic researchers, clinicians and drug developers working in the fields of metabolism and stem cells to provide a forum for interdisciplinary discussion and cross-collaboration amongst experts in these rapidly advancing fields.

Symposium themes:

- Metabolic regulation of stem cell self renewal
- Niche influence on stem cell metabolism
- Energetics of stem cell flux
- Mitochondria regulation of stem cell fate
- Metabolites in stem cell epigenetics and reprogramming
- Altered stem cell metabolism in aging and disease

Organizers

Nikla Emambokus, Editor, *Cell Metabolism*

Sheila Chari, Scientific Editor, *Cell Stem Cell*

Emmanuelle Passegue, *University of California, San Francisco, USA*

Michael Teitell, *University of California, Los Angeles, USA*

www.cell-symposia-stem-cell-energetics.com

Biological Code Breaking

The new movie *The Imitation Game* (to be released after this article goes to press), which recounts the Nazi code-breaking triumph of Alan Turing, appears unlikely to devote much screen time to the mathematician's foundational contributions to biology. For cinematic reasons this is perhaps understandable, yet the occasion nevertheless provides a fitting opportunity to reflect on his seminal work "The Chemical Basis of Morphogenesis." In it, he proposed in mathematical terms how an initially homogeneous collection of cells might generate self-organizing morphological patterns. He made many theoretical predictions that were only much later confirmed, and his underlying conceptual framework—exploring the principles of biological self-organization—remains as timely as ever.



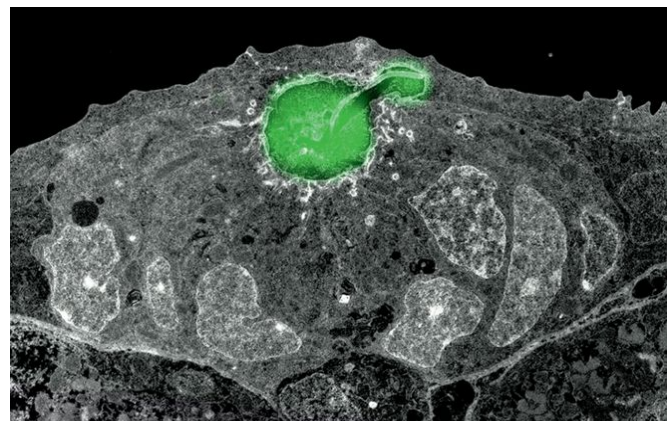
Digit patterning during mouse limb development. Image courtesy of J. Sharpe.

Unknown in Turing's day, there are now numerous beautifully worked-out examples of the phenomena he proposed—systems involving diffusible activators and inhibitors that interact to generate self-organized patterns. The most recent of these, elegantly uncovered by Raspopovic et al. (2014), describes a Turing network involved the formation of digits from developing limb buds. The authors make the observation that *Sox9*, a skeletal marker essential for digit formation, has a periodic pattern of expression in the limb bud at a very early stage, and based on the expectation that diffusible regulators in Turing-like processes would be either in phase or out of phase with *Sox9* expression, the authors identify pathways differentially activated in *Sox9*⁺ and *Sox9*[−] cells. This analysis implicates the morphogens BMP and WNT, whose involvement with *Sox9* is then explored by a combination of simulation and experimental manipulation. A compelling prediction from the modeling, which is borne out in pharmacological experiments in mouse limb bud cultures, is that combined inhibition of BMP and WNT, rather than simply inhibiting or boosting gene expression, is able to actively

reorganize the spatial pattern of digits, such that larger digits are formed within the same space, thus resulting in the formation of just four, three, or two digits instead of the usual five.

Reaction-diffusion systems are also found on a subcellular scale. The interaction between MinD and MinE proteins in *E. coli*, for instance, make the two proteins oscillate between cell poles, establishing time-averaged gradients that direct the site of cell division. Zieske and Schwille (2014) have recently succeeded in reconstituting these oscillations in a cell-free system (microengineered soft-polymer compartments) and show that these gradients facilitate the localization of FtsZ, a downstream mediator of cell division, to the center of these simulated cells. These chambers can be readily made in different shapes and sizes, thus permitting the systematic exploration of the role of cell geometry in forming intracellular protein gradients.

Perhaps nowhere has the theme of biological self-organization been more prominently on display as of late than in the organoid field. We are learning from these remarkable in vitro systems that individual cells harbor an extraordinary amount of intrinsic know-how for making complex tissues, a latent ability that is waiting to be unleashed. For example, the key to making gastric organoids, as newly reported by McCracken et al. (2014), is walking human pluripotent stem cells through the normal stages of gastric development (induction of definitive endoderm to primitive foregut to antral specification) with carefully-timed combinations of growth factors. With only this handful of external interventions, the cultured cells do the rest, recapitulating many of the complex features and cell types found in vivo, including the generation of cells expressing markers of stem and endocrine cells. When the ulcer-causing microbe *H. pylori* is introduced into the organoid cultures, an increase in epithelial cell proliferation is observed within 24 hr, illustrating the power of this (and other organoid models) for studying the initiating events in human disease.



Correlative light electron microscopy image of secreted green fluorescent protein concentrated in microlumen. Image courtesy of D. Gilmour.



Abstract
submission
deadline:

August 22
2014

Early bird
deadline:

October 10
2014

Join us in Belgium in December 2014 as we assemble some of the leading investigators examining various aspects of type 2 immunity in order to identify commonalities as well as contrasts of the cellular and molecular participants. This cross-fertilization will facilitate insights into this group of biological responses in health and disease.

Symposium themes:

- Role of innate immune cells in type 2 immunity
- Coordinating innate and adaptive type 2 inflammation
- Type 2 immunity in asthma and allergy and airway infection
- Type 2 immunity in microbial infection
- Role of commensal flora in influencing type 2 immunity

Keynote Speaker

Alan Sher, *NIAID, USA*

Visit our website to check out the full speaker line up

Organizers

Bruce Koppelman, Scientific Editor, *Immunity*

Ella Hinson, Scientific Editor, *Cell Host & Microbe*

Bart Lambrecht, *VIB and Ghent University, Belgium*

David Artis, *University of Pennsylvania, USA*

www.cell-symposia-type2-immunity.com

www.cell.com/symposia

Although organoids will continue to aid in understanding the conversations cells have with one another that confer self-organizing properties on these systems, equally enlightening insights will come from in vivo exploration, as evidenced by a recent study of the zebrafish lateral line primordium. Durdu et al. (2014) find that among the migrating cells of the primordium, the morphogen FGF selectively accumulates in microluminal structures and that signaling from this central lumen coordinates cell-cell communication to control the deposition and formation of mechanosensory organs. Given that many structures in vitro, such as organoids, and in vivo form lumina, this mode of signal concentration and cellular coordination could prove a widespread mode of tissue self-organization.

REFERENCES

- Durdu, S., Iskar, M., Revenu, C., Schieber, N., Kunze, A., Bork, P., Schwab, Y., and Gilmour, D. (2014). *Nature* 515, 120–124.
- McCracken, K.W., Catá, E.M., Crawford, C.M., Sinagoga, K.L., Schumacher, M., Rockich, B.E., Tsai, Y.H., Mayhew, C.N., Spence, J.R., Zavros, Y., and Wells, J.M. (2014). *Nature*. Published online October 29, 2014. <http://dx.doi.org/10.1038/nature13863>.
- Raspopovic, J., Marcon, L., Russo, L., and Sharpe, J. (2014). *Science* 345, 566–570.
- Zieske, K., and Schwille, P. (2014). *eLife*. <http://dx.doi.org/10.7554/eLife.03949>.

Robert P. Kruger

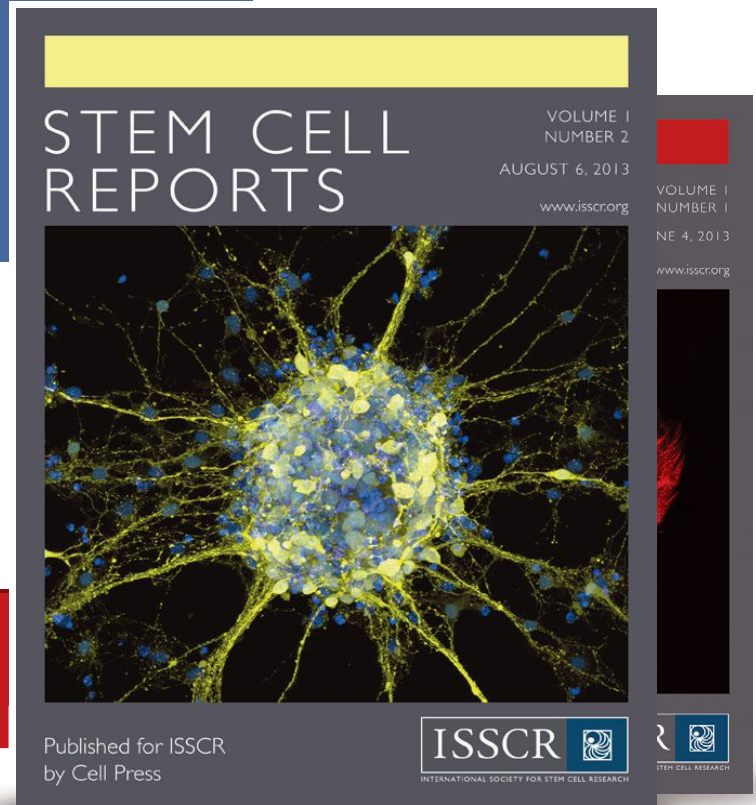
INTRODUCING STEM CELL REPORTS

EDITOR-IN-CHIEF

Christine Mummery, PhD,
Leiden University Medical Center

**CHECK OUT THE
LATEST ARTICLES ONLINE**

cell.com/stem-cell-reports



“ This open-access forum will
provide high visibility for your
stem cell research ”

Shinya Yamanaka, MD, PhD
ISSCR President and Nobel Laureate

A Place and a Grid in the Sun

Eric Kandel^{1,2,3,*}

¹Kavli Institute for Brain Science, Columbia University, 1051 Riverside Drive, New York, NY 10032, USA

²Howard Hughes Medical Institute, Columbia University, 1051 Riverside Drive, New York, NY 10032, USA

³Departments of Neuroscience, Biochemistry and Molecular Biophysics, and Psychiatry, Columbia University, 1051 Riverside Drive, New York, NY 10032, USA

*Correspondence: erk5@columbia.edu

<http://dx.doi.org/10.1016/j.cell.2014.11.033>

The 2014 Nobel Prize in Physiology or Medicine, awarded to John O'Keefe, May-Britt Moser, and Edvard I. Moser, recognizes the first deep-brain insights into a cognitive function. Their insights established a new view for how the brain represents spatial location.

This year the Nobel Prize Committee honored the discoveries, by John O'Keefe and by Edvard and May-Britt Moser, respectively, of the neural encoding of spatial location through place cells and grid cells. These groundbreaking contributions to cognitive neuroscience provided our first understanding of the role of the hippocampus and the entorhinal cortex in the brain's representation of space, which is likely to be critical for an animal's ability to form associative memories between a particular location in its environment and a particular sensory context. Their studies have also given us our first insights into a cognitive representation in higher brain regions that is concerned not simply with primary sensory or motor representation, but with a complex and abstract combination of modalities.

The Origins of the Problem, What Does the Hippocampus Do?

Classic studies by Brenda Milner and William Scoville in 1957 established the importance of the hippocampus for encoding declarative memory, the recall of information about people, places, objects, and events. Despite this advance, no one knew anything about the sensory signals that activated hippocampal neurons; in particular, no one thought the hippocampus might be important in the representation of space. As a result, it came as a great surprise when, in 1971, John O'Keefe and his student Jonathan Dostrovsky discovered that the pyramidal cells of the hippocampus encode not a single sensory modality—not touch, vision, taste, or smell—but something abstract, a representation, or

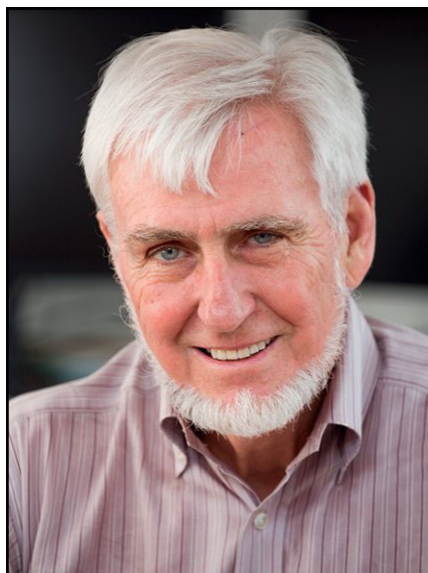
cognitive map, of space (O'Keefe and Dostrovsky, 1971, O'Keefe and Nadel, 1978).

O'Keefe and Dostrovsky discovered that various hippocampal neurons in the CA1 region of a rat's brain fired briefly when the animal assumed different positions in space. O'Keefe called these neurons "place cells" and the location that triggers the firing of each cell its "place field." By recording from many place cells at the same time, he further noticed that the firing of a given cell or set of cells forms a map indicating where in space the animal is located (O'Keefe, 1976; O'Keefe and Conway, 1978). When O'Keefe moved the rat to another space, the animal would form a new map using some of the same place cells as well as

some other place cells. When he brought the rat back to the original space, the initial cells would fire again, reforming the initial map. O'Keefe went on to demonstrate that the orientation of this map relative to an environment can be changed from trial to trial, and that this orientation is "remembered" even without spatial cues (O'Keefe and Speakman, 1987).

O'Keefe's second major insight was that the organization of the spatial map in the hippocampus is radically different from that of the sensory maps for touch and vision in the cortex. The maps for touch and vision are typically organized topographically, as Wade Marshall, Vernon Mountcastle, David Hubel, and Torsten Wiesel had discovered earlier; that is, neighboring cells in the cortex convey information about neighboring areas of the sensory periphery. The brain's spatial map is not organized topographically; thus, neighboring cells in the hippocampus do not convey information about neighboring positions in space. In fact, the spatial map is random with no topographic relation between neighboring cells and the part of external space they represent (O'Keefe and Nadel, 1978). This discovery, too, was totally unexpected and represents an entirely new view of how information is organized in the mammalian brain. It is generally considered the most important new finding in the sensory physiology of the mammalian brain since the Hubel-Wiesel finding of linear receptive fields in the primary visual cortex.

O'Keefe's third major contribution was theta phase precession—the finding that neurons in the hippocampus encode



John O'Keefe. Image courtesy of UCL.



May-Britt Moser. Image courtesy of Geir Mogen/NTNU.

location in space through both a rate code (firing frequency) and a temporal code (when the neurons fire). The hippocampus generates a coordinated pattern of activity as an animal explores its environment. This activity is produced by a network of neurons and gives rise to an extracellular electrical signal measured in an EEG called the theta rhythm, or oscillation, which ranges in frequency from 4 to 12 Hz. In 1993, O'Keefe and Michael Reece found that as an animal moves through a given neuron's place field, the timing of spike firing shifts progressively to earlier phases of the theta oscillation (O'Keefe and Recce, 1993). Because the place fields of different cells overlap, each place cell will fire at a different phase of the theta oscillation, enabling the brain to determine the animal's position with good precision (Huxter et al., 2003). This phenomenon of phase precession provides one of the few well-characterized examples in which the brain uses "temporal coding" that is not directly associated with the temporal properties of the stimulus.

Enter May-Britt and Edvard Moser

When May-Britt and Edvard Moser entered the field, they focused on analyzing how the neural circuitry of the hippocampus relates to spatial behavior. The hippocampus receives two types of

connections from the entorhinal cortex. One is via the direct pathway that carries information from the entorhinal cortex to area CA1 of the hippocampus. The other is via the indirect pathway from the entorhinal cortex via the dentate gyrus and the CA3 region of the hippocampus. They discovered that the brain's representation of space develops independently in areas CA3 and CA1 of the hippocampus. They then went on to show that the direct pathway from the entorhinal cortex to area CA1 is sufficient to maintain a spatial map: for the recognition of space; the indirect pathway is not required (Brun et al., 2002).

Having demonstrated that the spatial map is not intrinsic to the hippocampus and that the direct pathway is important, the Mosers then asked: how is the spatial map formed from this direct input? Is there a code for space in the entorhinal cortex, the major route of excitatory input to the hippocampus? In 2004 the Mosers discovered that, indeed, certain neurons in the medial part of the entorhinal cortex do encode an animal's location in space (Fyhn et al., 2004). However, unlike the place cells in the hippocampus, each of which encodes a unique place field, neurons in the entorhinal cortex represent space in a grid-like firing pattern (Hafting et al., 2005). Each cell fires not in a single location in space, but in multiple, evenly spaced locations. These locations form a periodic triangular, grid-like array, with clear regions of silence between the vertices of the triangles.

As the Mosers expanded the size of the environment the rat explored, they found that the grid-like firing field of each individual entorhinal neuron repeats itself across the entire space. Each cell imposes its pattern of firing on every environment the animal encounters. Thus, while place cells in the CA1 region provide a local description of space, grid cells provide a global description (Hafting et al., 2005; Fyhn et al., 2007). The finding of grid cells helped pinpoint the entorhinal cortex as a key hub in the brain network that enables us to find and remember our way.

The Mosers next searched for additional types of cells in the entorhinal cortex that might code for space. They found two: head direction cells (first described



Edvard I. Moser. Image courtesy of Geir Mogen/NTNU.

by Jim Ranck in the subiculum) that respond to the direction of the animal's head with respect to the environment and border cells that respond to the presence of a border or edge in the environment (Sargolini et al., 2006; Solstad et al., 2008). These two cell types generate a continuously updated representation of the position of the animal in space that can be used in any environment, irrespective of shape and landmarks.

These findings raised the question: Are the inputs from the grid cells in the entorhinal cortex the only ones that contribute to the formation of hippocampal place cells? The Mosers next went on to provide evidence that place cells are most likely generated by the convergence of signals from all three entorhinal cell types. The first evidence for this convergence emerged when the Mosers and John O'Keefe independently examined the causal connection between grid cells and place cells during development. They argued that if place cells are generated exclusively from grid cells, then grid cells should be active before, or at least simultaneously with, the emergence of place cell activity. However, recordings from rat pups revealed that this is not the case (Langston et al., 2010). At 2 weeks of age, when rats first leave the nest, a large proportion of hippocampal

cells already have place fields and therefore function as place cells, whereas grid cells exhibit only weak periodic fields. This suggests either that the weak spatial periodicity of the grid cells is sufficient to generate well-defined place fields in the hippocampus, or, more likely, that place cells are formed from the combined activity of grid cells and the two other types of entorhinal cells, particularly border cells, which exhibit some properties of adult cells at 2 weeks.

These results suggested that individual place cells receive important inputs from both grid cells and border cells, with grid cells providing information about distance based on the animal's motion and border cells providing the animal's position in relation to geometric boundaries. The strongest input seems to be from grid cells, which are several times more abundant than border cells, but under most circumstances the two classes of input are coherent and redundant: if one is absent, the other is sufficient to generate localized firing of place cells in the hippocampus.

Finally, since the hippocampus is known to send a reciprocal output to the entorhinal cortex, the Mosers examined how inactivation of the hippocampus influences entorhinal activity (Bonnevie et al., 2013). They found that silencing the hippocampus leads to a loss of grid cell periodicity and transforms grid cells into head direction cells. Thus, the encoding of space is not generated by a hierarchical, linear array of synaptic connections from the entorhinal cortex to the hippocampus, but rather depends on a recurrent loop of interconnections from the entorhinal cortex to the hippocampus and back to the entorhinal cortex.

The discoveries of grid cells by the Mosers and place cells by O'Keefe provide us with a new view of how the brain represents position in space. This spatial map, as I have emphasized, is a true cognitive function. There is no single modality, no single sensory organ, that determines location. Space is computed through a number of different sensory modalities—vestibular input, touch, vision, and smell—and it represents the coordinated totality of that input. In this context, grid cells are particularly interesting because they are

constructed globally from these sensory inputs.

The Work of O'Keefe and the Mosers Is Characterized by a Remarkable Intellectual Continuity

John O'Keefe received his Ph.D. from McGill University in 1967, when cognitive neuroscience was just emerging. The chairman of the psychology department at that time was Donald Hebb, who had studied with Karl Lashley and was exploring the problems of spatial orientation and spatial learning. Hebb went on to write "The Organization of Behavior: A Neuropsychological Theory," a classic text in which he argued that all behavior must be studied through the brain. Hebb, in turn, influenced the psychologist Brenda Milner, who, with the neurosurgeon Wilder Penfield, examined the hippocampus and its role in memory storage. As a postdoctoral fellow, O'Keefe worked with Patrick Wall at University College London. There, together with Dostrovsky, O'Keefe discovered place cells by analyzing the various factors that influence the firing properties of hippocampal neurons.

Together with Lynn Nadel, O'Keefe appreciated the historical importance of Edward Tolman's ideas. One of the earliest cognitive psychologists, Tolman argued in 1948 that rats running in a maze acquire "something like a field map of the environment" (Tolman 1948). Moreover, Tolman held that stimuli influence behavior not through stimulus-response connections, as Skinner and the behaviorists argued, but through the mediation of a cognitive map. Tolman thought of cognitive maps as broad, comprehensive maps or narrow strip maps confined to knowledge of specific groups.

O'Keefe's finding of a map was interesting enough in its own right, but of even greater interest was the fact that it is organized through a random population of cells. The idea of random organization of neural populations is now emerging as a broad theme in studies of higher regions of the cortex, for example the piriform cortex and the sense of smell (Stettler and Axel, 2009).

May-Britt and Edvard Moser were students of Per Andersen, an early pioneer in hippocampal electrophysiology, at

the University of Oslo. There, the Mosers met and married, and began their long-lasting successful personal and scientific collaboration. Indeed, with this award the Mosers became one of a select group of married couples, including Marie and Pierre Curie, to receive the Nobel Prize. With Per Andersen, the Mosers studied mechanisms of memory formation in the hippocampus in freely moving animals. In 1995–1996, they worked as postdoctoral fellows with Richard Morris, investigating the role of long-term potentiation in hippocampal memory. Then in 1996, they spent a brief but productive three month period with John O'Keefe, learning tetrode recording in the hippocampus. This experience started them on a quest that resulted in the remarkable discovery of grid cells, which linked the early classical work on the hippocampus to later studies of the hippocampus's input region in the entorhinal cortex. Thus, the Mosers are distinctive not only because of their discovery of entorhinal grid cells, border cells, and head direction cells—and the role these cells play in the formation of place cells—but also because of their important contribution to the physiology of the hippocampus (Leutgeb et al., 2007).

Such intellectual continuity represents science at its best, and this level of excellence is made even better by the fact that these three remarkable scientists are also extraordinarily fine human beings.

O'Keefe is currently Director of the Sainsbury Wellcome Center for Neural Circuits and Behavior at University College London. May-Britt Moser is Director of the Center for Neural Computation in Trondheim. Edvard I. Moser is Director of the Kavli Institute for Systems Neuroscience in Trondheim.

REFERENCES

- Bonnevie, T., Dunn, B., Fyhn, M., Hafting, T., Derdikman, D., Kubie, J.L., Roudi, Y., Moser, E.I., and Moser, M.B. (2013). *Nat. Neurosci.* 16, 309–317.
- Brun, V.H., Otnass, M.K., Molden, S., Steffenach, H.A., Witter, M.P., Moser, M.B., and Moser, E.I. (2002). *Science* 296, 2243–2246.
- Fyhn, M., Molden, S., Witter, M.P., Moser, E.I., and Moser, M.-B. (2004). *Science* 305, 1258–1264.

- Fyhn, M., Hafting, T., Treves, A., Moser, M.-B., and Moser, E.I. (2007). *Nature* 446, 190–194.
- Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., and Moser, E.I. (2005). *Nature* 436, 801–806.
- Huxter, J., Burgess, N., and O'Keefe, J. (2003). *Nature* 425, 828–832.
- Langston, R.F., Ainge, J.A., Couey, J.J., Canto, C.B., Bjerknes, T.L., Witter, M.P., Moser, E.I., and Moser, M.-B. (2010). *Science* 328, 1576–1580.
- Leutgeb, J.K., Leutgeb, S., Moser, M.B., and Moser, E.I. (2007). *Science* 315, 961–966.
- O'Keefe, J. (1976). *Exp. Neurol.* 51, 78–109.
- O'Keefe, J., and Conway, D.H. (1978). *Exp. Brain Res.* 31, 573–590.
- O'Keefe, J., and Dostrovsky, J. (1971). *Brain Res.* 34, 171–175.
- O'Keefe, J., and Nadel, L. (1978). *The Hippocampus as a Cognitive Map* (Oxford: Oxford University Press).
- O'Keefe, J., and Recce, M.L. (1993). *Hippocampus* 3, 317–330.
- O'Keefe, J., and Speakman, A. (1987). *Exp. Brain Res.* 68, 1–27.
- Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B.L., Witter, M.P., Moser, M.-B., and Moser, E.I. (2006). *Science* 312, 758–762.
- Solstad, T., Boccara, C.N., Kropff, E., Moser, M.-B., and Moser, E.I. (2008). *Science* 322, 1865–1868.
- Stettler, D.D., and Axel, R. (2009). *Neuron* 63, 854–864.
- Tolman, E.C. (1948). *Psychol. Rev.* 55, 189–208.

Better Imaging through Chemistry

Ernst H.K. Stelzer^{1,2,3,*}

¹Buchmann Institute for Molecular Life Sciences (BMLS)

²Cluster of Excellence (CEF-MC)

³Biowissenschaften, Fachbereich 15, Institut für Zellbiologie und Neurowissenschaften
Johann Wolfgang Goethe-Universität, 60438 Frankfurt am Main, Germany

*Correspondence: ernst.stelzer@physikalischebiologie.de

<http://dx.doi.org/10.1016/j.cell.2014.11.032>

The 2014 Nobel Prize in Chemistry has been awarded jointly to William E. Moerner, Stefan W. Hell, and Eric Betzig “for the development of super-resolved fluorescence microscopy.” I discuss the contributions made by this year’s awardees and how advances in understanding the behavior of fluorophores and research in light microscopy converged to allow the improved visualization of biological structures.

The recognition of the work that was granted in this year’s Nobel Prize in Chemistry highlights an important convergence of fields. William Moerner’s key studies on the characteristics of fluorophores and advances in the field of light microscopy by Stefan Hell and Eric Betzig have provided approaches that allow us to get a better picture of subcellular structures.

William Moerner has the distinction of being the first person to conceive of and perform experiments with single fluorophores (Moerner and Kador, 1989). Although Tomas E. Hirschfeld was able to systematically observe small assemblies of molecules about 13 years earlier, he was unable to visualize fewer than ~100 molecules (Hirschfeld, 1976). Moerner’s seminal 1997 paper showed that individual fluorophore molecules switch on and off, thereby sparking new ideas on how the photoactivation properties of fluorophores can be used to localize single molecules in a solid or even a fluid phase (Dickson et al., 1997). As such, Moerner’s studies lay the groundwork for the numerous localization microscopy approaches developed by Stefan Hell, Eric Betzig, and others.

Stefan Hell worked in my laboratory at EMBL as a postdoc during the early 1990s on 4Pi microscopy, an axial interference-based laser scanning confocal microscopy approach (Hell and Stelzer, 1992; Hell et al., 1994). By late 1993, Hell joined forces with a former collaborator of mine in Turku, Finland. While Steffen Lindek and I worked on confocal theta and 4Pi-theta fluorescence microscopy

(Stelzer and Lindek, 1994), which we later developed into light sheet-based fluorescence microscopy (Huisken et al., 2004), Hell and Jan Wichmann published their theoretical paper on stimulated emission depletion (STED) microscopy in 1994 (Hell and Wichmann, 1994). Although Hell explicitly received the Nobel Prize in Chemistry for STED, it should be noted that he systematically explored various schemes that deplete states in a fluorophore in his work on reversible saturable optical fluorescence transitions (RESOLFT).

Eric Betzig did much of his early work in scanning near-field optical microscopy (SNOM) and its applications to studies of

single fluorophores, addressing fundamental questions concerning the response of fluorophores to polarized light and the properties of the surfaces on which they were spread out. SNOM was regarded as one of the first techniques that could operate beyond the diffraction limit established in far-field light microscopy. He left Bell Labs and the scientific community in the late 1990s to work in his father’s company and returned in 2005 with a paper on optical lattice microscopy (Betzig, 2005). In 2006, Betzig and Harald Hess published a paper in *Science* describing how photoactivatable fluorophores could be used to measure the location of single molecules in a plane, with a precision in the tens of nanometers (Betzig et al., 2006). Using this photoactivated localization microscopy (PALM) approach, Betzig and Hess were able to generate superresolved images.

Now, a question that may have arisen in many people’s minds is why a Nobel Prize for superresolution microscopy has been awarded in chemistry and not in physics? The answer is quite simple: none of the nanoscopic instruments rely on novel physical principles, nor have they produced new physics-related insights. Rather, they rely on the properties of fluorophores. Let us now have a look at the physical limits of resolution, and then come back to the properties of fluorophores.

Resolution is limited by the physical properties of light, whose diffraction through a grating was well-known but was rephrased by Ernst Abbe in the



William E. Moerner. Image courtesy of L.A. Cicero/Stanford News.



Stefan W. Hell. Image courtesy of Deutscher Zukunftspreis - A. Pudenz.

context of microscopy (Abbe, 1873). Although light interference patterns are complex and are characterized by multiple orders, Abbe suggested that to get an accurate image of a specimen, it was sufficient to consider only undiffracted and first-order diffracted light. This means that the diffraction limit equation can be collapsed down into a relatively simple relationship as follows:

$$m \lambda = 2 d n \sin \alpha \quad \wedge \quad m = 1 \quad \Rightarrow$$

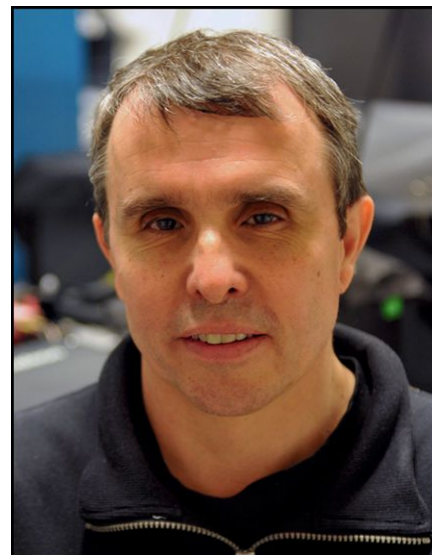
$$d = \frac{\lambda}{2n \sin \alpha},$$

where d relates to the spacing of the grating and is commonly regarded as the lower limit to the lateral resolution of an optical instrument, n is the refractive index of the material the light is passing through, and α is the angle into which the light beam is diffracted. However, the relationship is only correct for a uniform illumination and a very thin, borderless grating. It is not valid for a grid and is, in a sense, a relationship for a two-dimensional world. Even phrased for a grid and for our three-dimensional world, the equation will always refer to the limits of an image-forming device and not necessarily to the limits of a sampling device. Thus, Abbe's resolution is not about localization or precision, and of course Abbe did not consider scanning devices. It should be noted that although a popular

view among many scientists is that Abbe's formula limits the lateral resolution to $\lambda/2$, a lateral resolution of $\lambda/5$ is achievable in an optical transmission microscope with an annular condenser (Vainrub et al., 2006). Indeed, the lateral resolution of a far-field microscope can easily be pushed to beyond $\lambda/4$ within the current theoretical framework.

So how does this relate to superresolution microscopy? Basically, all of these light optical methods rely on localized fluorescence, namely the specific observation of a subset of the potential targets in a specimen. They either excite the fluorophores deterministically, as in STED microscopy, or statistically, as in PALM and STORM. Notably, neither method actually forms an image that can be observed directly through an ocular. Rather, they are sampling devices that collect and interpret the signal and then calculate the image. This is the key to how superresolution microscopy manages to surpass the resolution limits that apply to ordinary imaging devices: the boost in resolution comes from sampling methods. Indeed, Betzig and Hess were fully aware of the fact that they had not developed an image forming device, but a statistical sampling instrument that localizes the source of the fluorescence emission by analyzing multiple images: hence the "L" in "PALM." For the same reasons, Xiaowei Zhuang refers to "reconstruction microscopy" in her development of STORM (Rust et al., 2006) and Sam Hess to "localization" in his development of FPALM (Hess et al., 2006).

How exactly do these approaches work? The basic idea of STED is to confine the excitation of the fluorophores to a small area or volume using two laser beams. First, fluorophores are excited with a diffraction-limited light spot, similar to what is done using a confocal fluorescence microscope. The second laser beam has a doughnut shaped intensity distribution with an essentially zero intensity in its center. While the fluorophores are in the excited state, the second laser beam depletes all fluorophores apart from those in its center. The diameter of the central area is determined by the intensity of the second doughnut shaped laser, with higher intensity resulting in smaller area. The two laser beams sample the specimen, and since the relative posi-



Eric Betzig. Image courtesy of M. Staley.

tions of the intensity measurements are known, the fluorescence response can be used to calculate an image. The precision in the resultant image is defined by the precision with which a scanner moves the two laser beams across the specimen. An increase of the recording speed by parallelization while still maintaining the lateral resolution in infinitely thin specimens requires a minimal distance between any two excitation beams, which is, unsurprisingly, essentially determined by Abbe's diffraction limit. This requirement applies essentially to any sampling-based microscopy.

On the other hand, in PALM and STORM, an estimate is made of the location from which ideally a single fluorophore emits light. The basic idea of stochastic methods is to excite a fraction of the potentially excitable fluorophores, collect their emission with a camera, and then determine the central location of each fluorophore. After Eric Betzig started a collaboration with Harald Hess, he joined forces with Jennifer Lippincott-Schwartz and George Patterson, who had developed photoactivatable GFP (Patterson and Lippincott-Schwartz, 2002) and performed the first experiments in their laboratory at NIH. This was the missing piece in his quest to develop the localization microscopy he had in mind for a number of years (Betzig, 1995). In PALM, a fraction of the GFP molecules

are photoactivated at low light intensity to excite them in a manner similar to regular wide-field fluorescence microscopy. The emitted light is then collected by a camera and analyzed to determine the central locations of all fluorophores, with the entire process repeated until all fluorophores have been bleached. The compiled list of fluorophore locations is finally used to calculate an image.

It should be noted that excitation of fluorophores with very high intensities can also be used to generate higher harmonics, and hence frequencies that support the visualization of objects outside the diffraction limit of a linearly illuminated specimen, as pointed out by Rainer Heintzmann (Heintzmann et al., 2002). The late Mats Gustafsson was the first to take advantage of this effect in a nonlinear structured illumination microscope (SIM) (Gustafsson, 2005). In SIM, at least two coherent illumination beams interfere in a fluorophore-labeled specimen and usually create a sinusoidal interference pattern, which can be imaged but not necessarily resolved with a camera (Gustafsson, 2000). Three or five phases are recorded to generate a single image, with a resolution improved by a factor of two, at best, in a linear SIM. In a nonlinear mode, SIM provides a lateral resolution of about 50 nm, which is certainly in the same league as PALM, STORM, or STED microscopy, and represented an early step in the right direction.

There are a number of limitations in superresolution microscopy, some of which apply to fluorescence microscopy in general. For example, the number of fluorescence photons that can be collected from a fluorophore-labeled specimen is limited. However, improving the resolution by a factor of two reduces the volume from which the signal stems by a factor of eight, meaning only one eighth the number of fluorophores contributes to a particular pixel in the image. This means that the recording time has to be increased by a factor of 64 to retain the same signal-to-noise ratio, but it is by no means clear that this is possible with regular fluorophores. In addition, the excitation light is also absorbed by endogenous organic molecules, which are then degraded, leading to cellular defects that make live measurements difficult. More-

over, as life on earth is adapted to the solar constant, which is usually less than $1,000 \text{ W/m}^2$ on the ground, irradiance should not exceed 100 mW/cm^2 when biological specimens are followed in time and space. Another problem is that high wavelengths (around 700 nm) of light are used for excitation in superresolution microscopy. This may be unnecessary in some cases; for instance, when imaging “dead and flat” specimens, smaller labeling dyes and 380 nm UV light might be used to yield better and much brighter images.

While many groups have been working on the physical and technical aspects of STED microscopy, STORM, and PALM, a major challenge has been the application of these techniques to obtain new biological insights beyond what might be learned from electron microscopy, biochemistry, or other simpler forms of light microscopy (e.g., total internal reflection fluorescence microscopy). The group of Stefan Hell has not only developed RESOLFT and mastered all the problems one usually encounters when a new method is established from scratch, but they have also made a heroic effort in applying STED in a number of different biological applications. These include examining processes within neuronal synapses, the distribution of lipids, and the observation of live events in intact mouse brains. This work clearly outlines the potential of their methods in the life sciences. In addition, localization microscopy delivers single-molecule information about molecular distributions. STORM and its variant dSTORM (Heilemann et al., 2008), can even measure the absolute number of proteins present in a subcellular compartment (Ehmann et al., 2014), providing the insights into biological systems that are required for developing quantitative models of complex biological interactions.

Looking forward, a number of groups have recently combined the superresolution techniques with methods that are more suited for thick specimens, such as light-sheet-based fluorescence microscopy (Keller et al., 2008). Francesca Cella Zanacchi and Alberto Diaspro have also combined static LSFM with PALM (Cella Zanacchi et al., 2011) to observe multicellular spheroids. A particularly interesting development by Eric

Betzig (Chen et al., 2014) combines LSFM and the concept of coherent structured illumination (SIM) to reduce the light sheet thickness by replacing the traditional Gaussian beam with several interfering Bessel beams. This makes LSFM suitable for the observation of thin specimens. Two to three cell layers and the superficial volume segments of thicker specimens can be observed. The resolution is comparable to that of a confocal fluorescence microscope, while recording speed, number of frames and duration are heavily improved. Thus, PALM can be used in a dynamic fashion and is not restricted to flat specimens anymore.

As we await more biological insights, one important point to keep in mind is how recently superresolution microscopy was developed. Perhaps a useful point of comparison would be the 1986 award of the Nobel Prize in Physics in part to Gerd Binnig and Heinrich Rohrer for scanning tunneling microscopy (Binnig et al., 1982) (STM). STM and atomic force microscope (AFM) were the first instruments that produced images of single molecules spread out on a flat surface. I can recall my own amazement upon actually seeing the structure of a benzene ring and the twist of double stranded DNA. Rather than being inferred, they were clearly visible and resembled the structures known from decade-old textbooks. This amazement probably came in part from the fact that my tutors in theoretical physics had told me not to rely on pictures, but to draw my entire faith in the mathematical descriptions. In general, much of the early fascination for these techniques did not stem from novel information, but rather from the confirmation that they provided. However, thirty years later, the situation is entirely different. Both techniques have established themselves as the methods of choice in many disciplines, all over the world. There is every reason to believe that the same will be true of superresolution microscopy.

ACKNOWLEDGMENTS

Research in the Stelzer lab is supported by the Deutsche Forschungsgemeinschaft (DFG, EXC-115) and the Bundesministerium für Bildung und Forschung (BMBF, ProMEBS).

REFERENCES

- Abbe, E. (1873). *Archiv für Mikroskopische Anatomie* 9, 413–468.
- Betzig, E. (1995). *Opt. Lett.* 20, 237–239.
- Betzig, E. (2005). *Opt. Express* 13, 3021–3036.
- Betzig, E., Patterson, G.H., Sougrat, R., Lindwasser, O.W., Olenych, S., Bonifacino, J.S., Davidson, M.W., Lippincott-Schwartz, J., and Hess, H.F. (2006). *Science* 313, 1642–1645.
- Binnig, G., Rohrer, H., Gerber, Ch., and Weibel, E. (1982). *Appl. Phys. Lett.* 40, 178–180.
- Cella Zanacchi, F., Lavagnino, Z., Perrone Donnorso, M., Del Bue, A., Furia, L., Faretta, M., and Diaspro, A. (2011). *Nat. Methods* 8, 1047–1049.
- Chen, B.C., Legant, W.R., Wang, K., Shao, L., Milkie, D.E., Davidson, M.W., Janetopoulos, C., Wu, X.S., Hammer, J.A., 3rd, Liu, Z., et al. (2014). *Science* 346, 1257998.
- Dickson, R.M., Cubitt, A.B., Tsien, R.Y., and Moerner, W.E. (1997). *Nature* 388, 355–358.
- Ehmann, N., van de Linde, S., Alon, A., Ljaschenko, D., Keung, X.Z., Holm, T., Rings, A., DiAntonio, A., Hallermann, S., Ashery, U., et al. (2014). *Nat Commun* 5, 4650–4661.
- Gustafsson, M.G. (2000). *J. Microsc.* 198, 82–87.
- Gustafsson, M.G. (2005). *Proc. Natl. Acad. Sci. USA* 102, 13081–13086.
- Heilemann, M., van de Linde, S., Schüttelz, M., Kasper, R., Seefeldt, B., Mukherjee, A., Tinnefeld, P., and Sauer, M. (2008). *Angew. Chem. Int. Ed. Engl.* 47, 6172–6176.
- Heintzmann, R., Jovin, T.M., and Cremer, C. (2002). *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 19, 1599–1609.
- Hell, S., and Stelzer, E.H.K. (1992). *Opt. Commun.* 93, 277–282.
- Hell, S.W., and Wichmann, J. (1994). *Opt. Lett.* 19, 780–782.
- Hell, S.W., Lindek, S., Cremer, C., and Stelzer, E.H.K. (1994). *Appl. Phys. Lett.* 64, 1335–1337.
- Hess, S.T., Girirajan, T.P., and Mason, M.D. (2006). *Biophys. J.* 91, 4258–4272.
- Hirschfeld, T. (1976). *Appl. Opt.* 15, 2965–2966.
- Huisken, J., Swoger, J., Del Bene, F., Wittbrodt, J., and Stelzer, E.H.K. (2004). *Science* 305, 1007–1009.
- Keller, P.J., Schmidt, A.D., Wittbrodt, J., and Stelzer, E.H.K. (2008). *Science* 322, 1065–1069.
- Moerner, W.E., and Kador, L. (1989). *Phys. Rev. Lett.* 62, 2535–2538.
- Patterson, G.H., and Lippincott-Schwartz, J. (2002). *Science* 297, 1873–1877.
- Rust, M.J., Bates, M., and Zhuang, X. (2006). *Nat. Methods* 3, 793–795.
- Stelzer, E.H.K., and Lindek, S. (1994). *Opt. Commun.* 111, 536–547.
- Vainrub, A., Pustovsky, O., and Vodyanoy, V. (2006). *Opt. Lett.* 31, 2855–2857.

Speciation via Autoimmunity: A Dangerous Mix

Nitin Phadnis^{1,*} and Harmit S. Malik^{2,3,*}

¹Department of Biology, University of Utah, Salt Lake City, UT 84112, USA

²Division of Basic Sciences

³HHMI

Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA

*Correspondence: nitin.phadnis@utah.edu (N.P.), hsmalik@fhcrc.org (H.S.M.)

<http://dx.doi.org/10.1016/j.cell.2014.11.028>

In this issue of *Cell*, Chae et al. find that genomic “hot spots” encoding NLR plant immune receptor genes are recurrently responsible for hybrid necrosis, highlighting the role of host-pathogen evolutionary arms races in driving the evolution of hybrid incompatibilities.

The evolution of new species involves the establishment of reproductive isolating mechanisms such as hybrid sterility or hybrid inviability between previously interbreeding populations (Coyne and Orr, 2004). Understanding the molecular basis of hybrid incompatibilities—the deleterious genetic interactions that are responsible for hybrid defects—is a profound problem in biology. Two questions are of particular interest in this regard. First, are particular genes or molecular pathways prone to playing a recurrent role in the manifestation of hybrid incompatibilities? Second, what are the biological forces that drive the emergence and spread of incompatible alleles in populations?

Studying hybrid incompatibilities within species provides a powerful approach to study speciation; the same hybrid incompatibilities that segregate within species may also provide the raw material for the establishment of reproductive isolation between species. For instance, hybrid necrosis is a commonly observed defect in many inter- and intraspecific crosses in plants. In this issue of *Cell*, Chae et al. provide a detailed genetic analysis of hybrid necrosis between strains of *Arabidopsis thaliana* collected from diverse geographical locations (Chae et al., 2014). The scale of their analysis is staggering. Through thousands of crosses involving 80 completely sequenced strains of *A. thaliana* (Cao et al., 2011), the authors identify 142 cases of F1 hybrid necrosis. Of these, seven were picked for further genetic analyses in which the causal allele was likely to be present in multiple genetic

backgrounds, as evidenced by the similarity of F1 hybrid phenotypes produced in crosses of one parent with several other parental backgrounds. An analysis of F2 offspring using genotyping by sequencing identified seven new hybrid necrosis loci, labeled *DM3* to *DM9* (*DM* stands for *Dangerous Mix* [Alcázar et al., 2009; Bomblies et al., 2007]). Intriguingly, nearly all of the *DM* loci encode plant immune NLR (nucleotide-binding domain and leucine rich) proteins (Spoel and Dong, 2012).

Many of the F1 incompatibilities underlying hybrid necrosis involve pairwise interactions between distinct NLR loci, in accordance with a common portrayal of the Bateson-Dobzhansky-Muller (BDM) model for hybrid incompatibilities (Figure 1A). Under this model, new alleles at separate loci can arise and become fixed in populations because they are compatible with the genetic backgrounds in which they arose. However, these new alleles cause problems in hybrids when they do not function properly together. Surprisingly, *DM8* and *DM9* involve deleterious heterozygote interactions at the same genetic locus. Such single locus hybrid incompatibilities are rare and have generally been thought to be unlikely because a new incompatible allele that is sufficient to cause hybrid dysfunction must necessarily originate in a heterozygous state and therefore be instantaneously deleterious. Examples include speciation between dextral and sinistral versions of snails (Orr, 1991; Ueshima and Asami, 2003). Another way out of this conundrum is the sequential fixation

of new alleles at the same locus, as appears to be the case with the *DM8*- and *DM9*-incompatible alleles (Figure 1B).

The identification of the incompatible NLR alleles of *DM8* and *DM9* also suggests a specific biochemical possibility to explain still poorly understood aspects of NLR protein activation. For instance, pioneering work studying mechanisms of plant NLR protein activation has revealed that they can be activated either directly by specific pathogen effectors (“non-self”) or by effector-mediated modifications of other “signaling hub” host proteins (“modified self”) such as the RPM1-INTERACTING PROTEIN 4 (RIN4) (Maekawa et al., 2011; Spoel and Dong, 2012) (Figure 1C). Yet, what molecularly activates plant NLR proteins is still somewhat mysterious. The fact that *DM* alleles can cause hybrid necrosis with such high penetrance must imply that this combination of NLR proteins leads to NLR activation. This could be because each NLR protein is incompatible with a variant of a modified self protein (e.g., RIM4) from the other genome, resulting in cross-activation (Spoel and Dong, 2012). However, if this were the case, the other locus should be genetically identifiable as a hybrid incompatibility locus. Alternatively, the direct interaction of these incompatible NLR proteins with each other might directly activate the host necrosis response. Under this scenario, an NLR protein might activate by recognizing an incompatible NLR partner as a modified self protein (Figure 1D). In this regard, it is probably not a coincidence that Chae et al. find the *DM2* locus to be involved

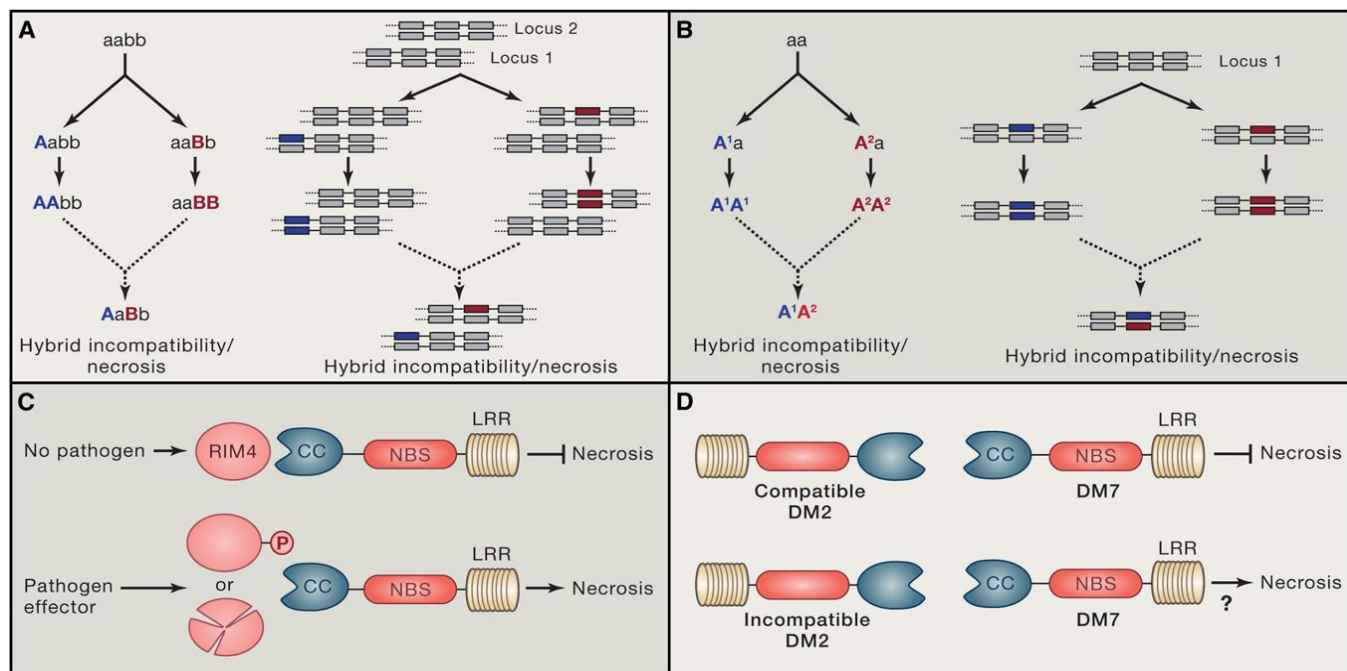


Figure 1. Hybrid Incompatibilities in NLR Genes Lead to Necrosis in *Arabidopsis*

(A) Consistent with the two-locus version of the Bateson-Dobzhansky-Muller (BDM) model for the evolution of hybrid incompatibilities, Chae et al. show that incompatible alleles could arise on different NLR clusters to cause incompatibility and hybrid necrosis.
 (B) *DM8* and *DM9* represent NLR alleles that each arose in isolation with compatible ancestral DM versions but become incompatible in the novel, untested configuration, consistent with the much rarer single-gene model of hybrid incompatibility.
 (C) Plant NLR genes frequently activate to trigger immunity (e.g., necrosis) upon sensing “self” proteins that have been modified by pathogen effectors i.e., “modified self” proteins. For instance, pathogen effector-mediated phosphorylation or fragmentation of the host protein RIM4 leads to specific activation of certain NLRs.
 (D) We speculate that the incompatibility between NLR proteins might arise because allelic differences in the *DM2* locus (compatible versus incompatible alleles) are interpreted as modified self proteins, triggering NLR activation and therefore hybrid necrosis.

in at least five out of nine known cases of F1 hybrid necrosis (Chae et al., 2014). *DM2* could represent an example of a signaling hub protein in which incompatible NLR proteins interpret allelic differences as modified self instead. *DM2* may be especially susceptible to triggering hybrid necrosis because its enzymatic activity may be directly coupled to downstream signaling events that induce cell death. Study of the biochemical differences between the hybrid necrosis risk and nonrisk alleles of NLR genes is likely to reveal insights into what activates NLR proteins and what keeps them in check.

Hybrid incompatibilities cause severely deleterious fitness consequences; the incompatible alleles are certainly not selected for these properties. Indeed, one might expect alleles with a propensity to be incompatible to have a significant selective cost within populations. Instead, hybrid incompatibilities are likely an accidental consequence of the evolution of

these genes for other reasons. Biological phenomena that drive the rapid and recurrent evolution of genes, such as intragenomic conflict (Phadnis and Orr, 2009) or host-pathogen arms races, provide strong candidates for the engine of speciation. The increased genetic repertoire of NLR genes in plants (150 in *Arabidopsis* and 450 in rice) as compared to vertebrates (~20) (Maekawa et al., 2011), together with their rapid evolution, might simply increase the odds of incompatible combinations, or “a dangerous mix.” This might explain why hybrid incompatibility due to NLR genes is so rampant in plants.

Crop breeders sometimes actively select hybrid necrosis genes to derive strains that are resistant to particular pathogens. It is almost certain that natural selection has done this for far longer, in a “tit-for-tat” between plant NLR proteins and plant pathogen effectors. Although this rapid adaption may confer pathogen resistance, Chae et al. find that such

immunity might recurrently levy an accidental but high cost in occasionally producing incompatible combinations that cause autoimmunity and hybrid necrosis (Chae et al., 2014), sowing the seeds of speciation.

ACKNOWLEDGMENTS

We are supported by start-up funds and a Mario Capecchi endowed chair in Biology from the University of Utah (N.P.) and by grants from HHMI, the Mathers Foundation, and NIH R01 GM074108 (H.S.M.). We thank Maulik Patel and Rick McLaughlin for comments.

REFERENCES

- Alcázar, R., García, A.V., Parker, J.E., and Raymond, M. (2009). Proc. Natl. Acad. Sci. USA 106, 334–339.
- Bomblies, K., Lempe, J., Eppe, P., Warthmann, N., Lanz, C., Dangl, J.L., and Weigel, D. (2007). PLoS Biol. 5, e236.
- Cao, J., Schneeberger, K., Ossowski, S., Günther, T., Bender, S., Fitz, J., Koenig, D., Lanz, C., Stegle,

O., Lippert, C., et al. (2011). *Nat. Genet.* 43, 956–963.

Chae, E., Bomblies, K., Kim, S.-T., Karelina, D., Zaidem, M., Ossowski, S., Martin-Pizarro, C., Laitinen, R.A.E., Rowan, B.A., Tenenboim, H., et al. (2014). *Cell* 159, this issue, 1341–1351.

Coyne, J.A., and Orr, H.A. (2004). *Speciation* (Sunderland, Mass.: Sinauer Associates).

Maekawa, T., Kufer, T.A., and Schulze-Lefert, P. (2011). *Nat. Immunol.* 12, 817–826.

Orr, H.A. (1991). *Evolution* 45, 764.

Phadnis, N., and Orr, H.A. (2009). *Science* 323, 376–379.

Spoel, S.H., and Dong, X. (2012). *Nat. Rev. Immunol.* 12, 89–100.

Ueshima, R., and Asami, T. (2003). *Nature* 425, 679.

A Split Personality for Nucleosomes

Daniel J. McKay¹ and Jason D. Lieb^{2,*}

¹Department of Biology, Department of Genetics, Integrative Program for Biological and Genome Sciences, The University of North Carolina at Chapel Hill, 250 Bell Tower Drive, Chapel Hill, NC 27599, USA

²Department of Human Genetics, The University of Chicago, 920 East 58th Street, Chicago, IL, 60637, USA

*Correspondence: jdlieb@uchicago.edu

<http://dx.doi.org/10.1016/j.cell.2014.11.030>

A high-resolution look at where histones touch DNA reveals a surprisingly intricate, dynamic, and modular nucleosome. Three advances in the study by Rhee et al. include unexpected interactions between the H3 tail and linker DNA, new evidence for existence of subnucleosomal particles, and asymmetric patterns of histone modification within a single nucleosome that correspond to the direction of transcription.

In eukaryotic cells, DNA is packaged into nucleosomes. A single nucleosome consists of a protein spool made of histones, wrapped by DNA. In addition to packaging DNA, nucleosomes also compete with other DNA-binding proteins and thereby influence access to the regulatory information that controls DNA-dependent processes such as transcription, replication, and DNA repair. In this issue of *Cell*, Pugh and colleagues (Rhee et al., 2014) apply a high-resolution mapping approach called ChIP-exo in yeast to examine the genome-wide position and organization of the individual histones that comprise nucleosomes. Their findings reveal surprisingly complex nucleosome substructures and dynamics that immediately bring to light an exciting set of new questions for the field, while at the same time evoking early models of the nucleosome (Weintraub et al., 1976).

Some background is required to set the stage for the three major advances derived from the results. The traditionally defined nucleosome core consists of an octamer of histone proteins, around which ~147 bp of DNA is wrapped. This octamer

is composed of two copies each of the histones H2A, H2B, H3, and H4. More specifically, dimers of H3 and H4 interact to form a tetramer, which is flanked on each side by a dimer of H2A and H2B. Pugh and colleagues used ChIP-exo to determine the precise location of individual histone proteins across the yeast genome. ChIP-exo is a modified version of conventional chromatin immunoprecipitation (ChIP) that provides high-resolution identification of binding sites for proteins that interact with DNA. Like ChIP, the first step in ChIP-exo is to covalently crosslink proteins to DNA with formaldehyde. After sonication to shear the chromatin into smaller fragments and immunoprecipitation with antibodies that recognize the protein of interest, ChIP-exo then uses lambda exonuclease to digest DNA strands in the 5' to 3' direction. Digestion is blocked when the exonuclease reaches a protein-DNA crosslink. After high-throughput sequencing, pairs of 5' ends on the forward and reverse strands (exonuclease stop points) thus represent the boundaries of a given protein-DNA interaction. ChIP-exo has previously been used to map binding sites for

sequence-specific transcription factors (Rhee and Pugh, 2011), preinitiation complexes (Rhee and Pugh, 2012), and chromatin remodelers (Yen et al., 2012; Yen et al., 2013).

The first intriguing result of Rhee et al. (2014) concerns the amino-terminal tail of histone H3, which is heavily decorated with posttranslational modifications and has important regulatory functions. ChIP-exo results for H2B and H4 histones identified crosslinking points that closely correspond to the genomic locations expected from the crystal structure (Luger et al., 1997). On the other hand, ChIP-exo results for histone H3 showed an unexpected crosslinking pattern. In the crystal structure, most of the amino acids comprising H3 reside at the nucleosome midpoint (called the “dyad”), where they contribute substantially to DNA interactions (Luger et al., 1997). However, the predominant H3-DNA interaction determined by ChIP-exo was located in the linker DNA that separates adjacent nucleosomes, not at the nucleosome dyad. The authors speculated that this interaction may be mediated through the N-terminal tail of histone H3, and then tested their

hypothesis by performing ChIP-exo in a yeast strain lacking the first 28 amino acids of H3. While the cross-linking signal between H3 and linker DNA diminished in the H3 Δ 1-28 strain, a noticeable signal remained that was still greater than the H3 signal at the dyad. Since approximately ten amino acids of the H3 tail are still present in the H3 Δ 1-28 strain, the source of the residual H3-linker signal remains unresolved. The significance of this interaction has not yet been tested, but it is notable due to the known regulatory functions of the H3 tail and because interactions between histone tails and linker DNA may be important for higher-order chromatin organization.

The increase in resolution afforded by ChIP-exo allowed the authors to make a second important finding: new evidence for semi-independent behavior of each half of the nucleosome and for the existence of subnucleosomal particles. The nucleosome crystal structure shows that the histones are roughly symmetrical about the dyad axis, with one H2A/H2B dimer and one H3/H4 dimer located on each half of the nucleosome. Therefore, one might expect that the histone occupancies on one half of the dyad would be tightly correlated with the histone occupancies on the other half of the dyad. While histone occupancies across the nucleosome are indeed correlated, the authors found that histone occupancies on a given half of the dyad correlated much more closely with each other than with the histone occupancies on the opposite half of the dyad. H2B in particular had pronounced differences in occupancy across the two halves of the nucleosome. These differences in histone occupancies could be explained by the presence of subnucleosomal-sized particles consisting of hexasomes (nucleosomes that lack one H2A/H2B dimer) and half-nucleosomes (nucleosomes containing one H2A/H2B dimer and one H3/H4 dimer). Multiple groups have provided evidence for subnucleosomal-sized particles previously (Floer et al., 2010; Henikoff et al., 2011; Kent et al.,

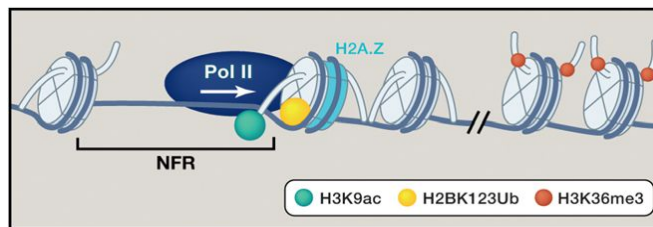


Figure 1. ChIP-Exo Reveals Asymmetric Features of Nucleosomes at Gene Promoters.

The histone variant H2A.Z is enriched at the +1 nucleosome, and is preferentially localized to the NFR-distal side of +1 nucleosomes with high turnover. At highly expressed genes, histone modifications linked to transcription (H2BK123ub and H3K9ac) are asymmetrically localized on the side of the +1 nucleosome closest to the NFR. H3K9ac is also associated with interactions between the H3 tail and DNA in the NFR. Away from promoters, histone H3 tails interact with linker DNA, except when H3K36 is methylated.

2011). These subnucleosomal particles were interpreted to be partially unwound nucleosomes, “fragile” nucleosomes, and nonhistone transcription factor complexes. The ChIP-exo data presented in Rhee et al. is consistent with previous findings and extends them by providing strong evidence that subnucleosomal-sized particles smaller than hexasomes can consist of half-nucleosomes rather than H3/H4 tetramers.

The third important observation involves asymmetries in the posttranslational modifications and histone variants within nucleosomes. Since a given gene is transcribed predominantly in a single direction, the histone modifications and variants that contribute to different steps of the transcription cycle are typically deposited with a distinct polarity along the length of the gene. For example, H3K4me3 is primarily found at gene promoters, whereas H3K36me3 is enriched toward the 3' end of genes. The authors asked if there was transcription-correlated polarity at the level of individual histone proteins within a nucleosome by performing ChIP-exo using antibodies that recognize H3K9ac, H2Bub, and the histone variant H2A.Z. Remarkably, the histone H3 crosslinks that remained following H3K9ac ChIP were enriched primarily on the NFR-proximal half of the +1 nucleosome. The +1 nucleosome is the first nucleosome in a gene immediately downstream of the transcriptional start site and is typically preceded by a nucleosome free region (NFR). ChIP-exo with antibodies to all forms of H3 revealed no crosslinking in the NFR, suggesting

that the H3K9ac interaction occurred only in a small subset of the population or was very transient. Further investigation revealed that this signal was only found at highly transcribed genes. H2Bub also showed a similar pattern of occupancy on the NFR-proximal half of the +1 nucleosome of highly transcribed genes. The histone variant H2A.Z was also incorporated into the +1 nucleosome; however, it was asymmetrically localized to the other side of the nucleosome, on the NFR-distal

half (Figure 1). This contrasts with previous ChIP-exo data in which H2A.Z showed symmetrical localization on the +1 nucleosome (Yen et al., 2013). Interestingly, the chromatin remodeler that deposits H2A.Z into chromatin (SWR-C) was reported to localize asymmetrically to the opposite (NFR-proximal) side of the +1 nucleosome. Therefore, how H2A.Z becomes enriched on the NFR-distal half of the +1 nucleosome remains to be determined. It is possible that an asymmetry exists in the proposed cycle of H2A.Z deposition and removal by chromatin remodelers (Yen et al., 2013).

The three principal findings described by Rhee et al., namely the interactions between histone H3 and linker DNA, evidence for existence of subnucleosomal particles, and the asymmetric patterns of histone modifications and variants, add to the evidence for a complex and dynamic nucleosome. Pugh and colleagues have discovered important details of nucleosome structure that will inform future investigations into the mechanisms by which histone organization, modification, and stability contribute to transcription and other DNA-dependent processes.

REFERENCES

- Floer, M., Wang, X., Prabhu, V., Berrozpe, G., Narayan, S., Spagna, D., Alvarez, D., Kendall, J., Krasnitz, A., Stepansky, A., et al. (2010). *Cell* 141, 407–418.
- Henikoff, J.G., Belsky, J.A., Krassovsky, K., MacAlpine, D.M., and Henikoff, S. (2011). *Proc. Natl. Acad. Sci. USA* 108, 18318–18323.

Kent, N.A., Adams, S., Moorhouse, A., and Paszkiewicz, K. (2011). *Nucleic Acids Res.* 39, e26.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). *Nature* 389, 251–260.

Rhee, H.S., and Pugh, B.F. (2011). *Cell* 147, 1408–1419.

Rhee, H.S., and Pugh, B.F. (2012). *Nature* 483, 295–301.

Rhee, H.S., Bataille, A.R., Zhang, L., and Pugh, B.F. (2014). *Cell* 159, this issue, 1377–1388.

Weintraub, H., Worcel, A., and Alberts, B. (1976). *Cell* 9, 409–417.

Yen, K., Vinayachandran, V., Batta, K., Koerber, R.T., and Pugh, B.F. (2012). *Cell* 149, 1461–1473.

Yen, K., Vinayachandran, V., and Pugh, B.F. (2013). *Cell* 154, 1246–1256.

How TriC Folds Tricky Proteins

Anastasia Zhuravleva^{1,*} and Sheena E. Radford^{1,*}

¹Astbury Centre for Structural Molecular Biology, School of Molecular and Cellular Biology, University of Leeds, Leeds LS2 9JT, UK

*Correspondence: a.zhuravleva@leeds.ac.uk (A.Z.), s.e.radford@leeds.ac.uk (S.E.R.)

<http://dx.doi.org/10.1016/j.cell.2014.11.029>

How chaperonins orchestrate the successful folding of even the most elaborate of proteins is a question of central importance. In two recent studies in *Cell* by Joachimiak et al. and Freund et al., a new class of TRiC substrate is identified, and how the chaperonin exploits its different subunits to extend its substrate repertoire and direct productive folding is revealed.

For proper functioning, newly synthesized proteins must be correctly folded. This can be difficult to achieve, especially for large proteins with complex topologies. Misfolded proteins are not only inactive but can be toxic, creating a devastating imbalance of protein synthesis and folding that has been linked to many devastating diseases (Kim et al., 2013). Molecular chaperones interact with unfolded and partially folded proteins to facilitate folding and prevent misfolding and aggregation. To perform these functions, ATP-driven molecular chaperones, such as Hsp70s, Hsp90s, and the Hsp60 chaperonins, use the energy of ATP to control substrate binding and release and to promote correct folding (Kim et al., 2013).

Chaperonins are complex allosteric machines. They consist of two stacked rings of seven or more identical, or homologous, subunits that form a barrel-like structure used to encapsulate the folding substrate protein (Figure 1). The most well-studied group I chaperonin, bacterial GroEL, is formed from two rings, each with seven identical 60 kDa subunits. This homo-oligomeric chaperonin interacts with its substrate proteins predominantly via hydrophobic interactions (Figure 1, left). By contrast, the eukaryotic group II chaperonin, TCP-1 ring complex

(TRiC), is a hetero-oligomeric chaperonin, and it recognizes its substrates via hydrophobic, electrostatic, and/or polar motifs (Dunn et al., 2001; Kalisman et al., 2013). The increased complexity of the hetero-oligomeric ring allows TRiC to promote folding of a very broad range of protein substrates. Indeed, about 5%–10% of all newly synthesized proteins require TRiC to fold (Yam et al., 2008). TRiC has also been shown to inhibit the aggregation of huntingtin, interacting with the tips of polyQ-containing fibrils, as well as smaller oligomers (Shahmoradian et al., 2013).

In this issue of *Cell*, Freund et al. (2014) report the discovery of a new TRiC substrate—the telomerase protein TCAB1—which is essential for trafficking of telomerase and small Cajal body RNAs required for telomere maintenance during cell division (Venteicher and Artandi, 2009). The authors performed a genome-wide RNA fluorescence in situ hybridization (FISH)-based siRNA screen for genes required for Cajal bodies' localization of a key telomerase enzyme, the telomerase RNA component (TERC), and telomerase protein TCAB1. Surprisingly, in addition to known telomerase assembly factors, the authors found that several TRiC subunits are required for TERC and TCAB1 locali-

zation in Cajal bodies. Depletion of TRiC results in a loss of TCAB1, mislocalization of telomerase and Cajal body RNAs, and failure of telomere elongation. TRiC, it turns out, is essential for TCAB1 folding. The results explain why mutations in TCAB1 can lead to severe diseases and suggest that a larger range of protein substrates than considered hitherto may require TRiC to fold.

In a second recent study in *Cell*, Joachimiak et al. (2014) shed exciting new light on the structural mechanism of substrate recognition by TRiC and how TRiC is able to fold its broad range of protein substrates (Yam et al., 2008, Shahmoradian et al., 2013, Freund et al., 2014). Each ring of TRiC consists of eight homologous subunits (CCT1–CCT8) (Figure 1, right-hand, top), with the majority of the sequence variations between TRiC's subunits being found in their apical domains (Dunn et al., 2001). Like its GroEL homolog, substrates bind to the apical domains of TRiC, and it has been suggested previously that the sequence variations in these domains are important for substrate recognition (Dunn et al., 2001; Kalisman et al., 2013). However, how TRiC binds its broad repertoire of substrates and promotes their correct folding remained elusive.

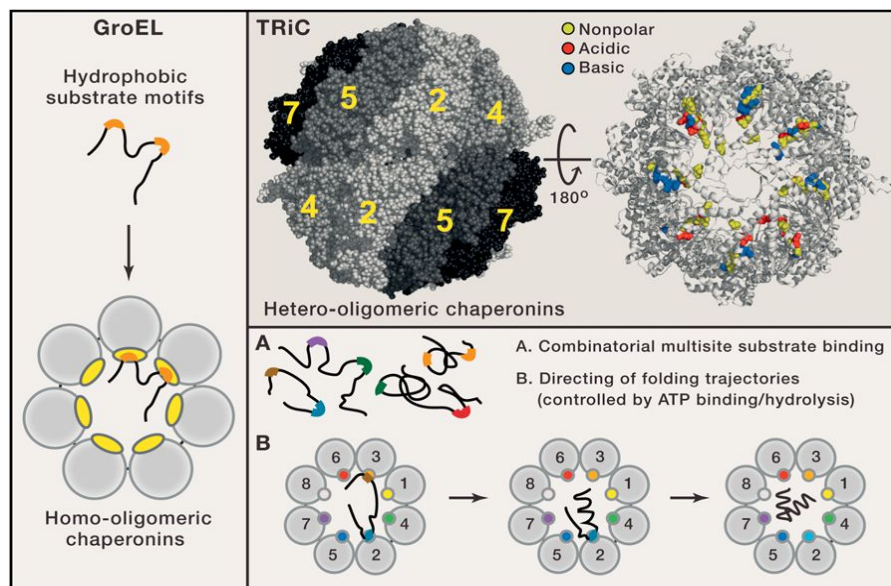


Figure 1. Substrate Recognition and Binding to Group I and Group II Chaperonins

(Left) Schematic model for substrate interactions with the group I chaperonin GroEL via hydrophobic motifs, which are identical for all seven subunits in each of its two rings. (Top right) X-ray structure of full-length hetero-oligomeric TRiC, which has eight different subunits in each of its two rings (PDB: 4D8R); amino acid types in the substrate-binding sites for individual apical domains are highlighted as colored spheres: nonpolar residues (yellow), polar (gray), acidic (red), and basic (blue). (Bottom right) Common rules for substrate binding by TRiC that allow folding of its different substrates: (A) combinatorial multisite substrate binding and (B) proposed model for how binding to different subunits may direct folding via sequential release of different recognition motifs. Different colors highlight unique substrate binding motifs for different TRiC subunits.

To address this question and to obtain a structural model for how TRiC binds its substrates, Joachimiak and colleagues used nuclear magnetic resonance (NMR) and modeling to determine the substrate-binding interface between the isolated apical domain of the TRiC subunit, CCT3, with the 54 residue HIV protein, p6. Alanine substitutions on the substrate-binding interface derived from this structural model revealed that nonpolar, polar, and charged residues contribute to the substrate binding kinetics for CCT3. Next, the authors explored whether other TRiC subunits share the same binding site. An extension of their analysis to the apical domain of CCT1 and its substrate (the so-called box 1 from Hippel Lindau tumor suppressor [VHL] [Spiess et al., 2006]) revealed that this substrate binds CCT1 in precisely the same region that CCT3 binds p6. In a similar vein, chemical crosslinking-mass spectrometry (XL-MS) was used to show that the TRiC substrates, tubulin and the HIV protein Gag, form multivalent contacts with different TRiC subunits

(CCT2, CCT6, and CCT7) using similar substrate-chaperonin interfaces to those identified for the isolated apical domains of CCT3 and CCT1 using NMR. Interestingly, the authors show how intrinsic flexibility of the substrate-binding site allows different substrates to bind in different configurations to the same apical domain, whereas the unique substrate-binding motifs in the different TRiC subunits enable different substrates that share no sequence similarity to bind (Figure 1, right-hand, lower). Asymmetric ATP binding to the TRiC ring (Reissmann et al., 2012) provides an additional level of complexity, which enables TRiC to release different regions of a protein substrate sequentially during its allosteric cycle. Such a mechanism would allow TRiC to orchestrate folding by controlled release of different regions of the substrate protein (which can then fold), whereas other regions remain bound to the TRiC ring. Together, the results reveal fascinating new insights into how a single chaperonin is not only able to fold an array of different protein sequences but also

how the route of folding may be manipulated by utilizing the different properties of individual subunits within the chaperonin ring (Figure 1).

Although TRiC-substrate interactions have come into clear focus through these exciting studies, several questions remain. To understand precisely how TRiC promotes folding of its different substrates, detailed structural information is needed to provide direct evidence for the appealing model proposed invoking directed folding via controlled substrate release by the chaperonin. Moreover, how TRiC binding is able to steer folding along productive routes and how the chaperonin is able to “choose” the right folding path for its different protein substrates remain a mystery. Discovery of new classes of TRiC substrates and further insights into ATP- and substrate-induced allosteric conformational changes within, and between, TRiC subunits for different TRiC substrates will be needed to answer these questions. Nonetheless, it is clear that nature has evolved a clever machine in TRiC that enables a single chaperonin to fold some of the trickiest of protein folds.

REFERENCES

- Dunn, A.Y., Melville, M.W., and Frydman, J. (2001). *J. Struct. Biol.* 135, 176–184.
- Freund, A., Zhong, F.L., Venteicher, A.S., Meng, Z., Veenstra, T.D., Frydman, J., and Artandi, S.E. (2014). *Cell* 159, this issue, 1389–1403.
- Joachimiak, L.A., Walzthoeni, T., Liu, C., Aebersold, R., and Frydman, J. (2014). *Cell* 159, 1042–1055.
- Kalisman, N., Schröder, G.F., and Levitt, M. (2013). *Structure* 21, 540–549.
- Kim, Y.E., Hipp, M.S., Bracher, A., Hayer-Hartl, M., and Hartl, F.U. (2013). *Annu. Rev. Biochem.* 82, 323–355.
- Reissmann, S., Joachimiak, L.A., Chen, B., Meyer, A.S., Nguyen, A., and Frydman, J. (2012). *Cell Rep.* 2, 866–877.
- Shahmoradian, S.H., Galaz-Montoya, J.G., Schmid, M.F., Cong, Y., Ma, B., Spiess, C., Frydman, J., Ludtke, S.J., and Chiu, W. (2013). *eLife* 2, e00710.
- Spiess, C., Miller, E.J., McClellan, A.J., and Frydman, J. (2006). *Mol. Cell* 24, 25–37.
- Venteicher, A.S., and Artandi, S.E. (2009). *Cell Cycle* 8, 1329–1331.
- Yam, A.Y., Xia, Y., Lin, H.T., Burlingame, A., Gerstein, M., and Frydman, J. (2008). *Nat. Struct. Mol. Biol.* 15, 1255–1262.

Metabolic Inflexibility: When Mitochondrial Indecision Leads to Metabolic Gridlock

Deborah M. Muoio^{1,*}

¹Sarah W. Stedman Nutrition and Metabolism Center, Duke Molecular Physiology Institute, Duke University Medical Center, Durham, NC 27710, USA

*Correspondence: muoio@duke.edu

<http://dx.doi.org/10.1016/j.cell.2014.11.034>

Normal energy metabolism is characterized by periodic shifts in glucose and fat oxidation, as the mitochondrial machinery responsible for carbon combustion switches freely between alternative fuels according to physiological and nutritional circumstances. These transitions in fuel choice are orchestrated by an intricate network of metabolic and cell signaling events that enable exquisite crosstalk and cooperation between competing substrates to maintain energy and glucose homeostasis. By contrast, obesity-related cardiometabolic diseases are increasingly recognized as disorders of metabolic inflexibility, in which nutrient overload and heightened substrate competition result in mitochondrial indecision, impaired fuel switching, and energy dysregulation. This Perspective offers a speculative view on the molecular origins and pathophysiological consequences of metabolic inflexibility.

Introduction

A staggering 68% of U.S. adults classify as obese or overweight. Increased adiposity is associated with insulin resistance, hypertension, hepatic steatosis, dyslipidemia, glucose intolerance, and hyperinsulinemia. Collectively known as the metabolic syndrome, this constellation of comorbidities raises the risk of developing cardiovascular disease and type 2 diabetes. In general, these are diseases of energy surplus, caused in large part by physical inactivity and overconsumption of calorically dense processed foods. Drug discovery efforts aimed at curtailing the epidemic spread of metabolic disease have focused heavily on mechanisms governing systemic glucose and lipid balance and the interplay between nutrient supply and insulin sensitivity. In most cases, the onset of insulin resistance comes early in disease development and plays a central role in the etiology of late-stage complications. Because insulin orchestrates systemic flux and disposal of glucose, fatty acids, and amino acids, resistance to the actions of the hormone gives rise to a metabolic storm of aberrant nutrient partitioning. Among the key features of this storm is an apparent stiffness in mitochondrial substrate selection, such that various organs and cell types fail to appropriately adjust fuel choice in response to nutritional circumstances. This phenomenon, dubbed “metabolic inflexibility,” has gained growing attention as a hallmark of cardiometabolic disease and a potential cause of cellular dysfunction. Thus, emerging evidence implies that metabolic health deteriorates as mitochondria lose their capacity to switch freely between alternative forms of carbon energy. This Perspective considers the physiological relevance of substrate choice and the molecular consequences of mitochondrial indecision.

Metabolic Flexibility and the Freedom of Choice

Mitochondria, the respiratory engines of the cell, consume oxygen to “burn” carbon intermediates derived from three principal

nutrients: fatty acids, glucose, and amino acids. These fuels can each be catabolized to acetyl-CoA, which serves as the universal substrate that feeds the tricarboxylic acid cycle (TCAC). Each turn of the TCAC releases carbons in the form of CO₂ while also generating reducing equivalents (NADH and FADH₂) that drive the electron transport chain (ETC) and oxidative phosphorylation (OXPHOS), a less powerful but more efficient and higher capacity ATP regenerating system than glycolysis. The ETC/OXPHOS system requires oxygen as the final electron acceptor, resulting in the production of water. Cellular rates of CO₂ production relative to oxygen consumption, or the respiratory quotient (RQ), fluctuate between 0.7 and 1.0 and provide an approximation of mitochondrial fuel use under typical conditions in which amino acids contribute only minimally as an oxidative substrate. A high RQ is indicative of glucose oxidation, whereas a low RQ reflects predominately fat oxidation.

Normal physiology is characterized by diurnal oscillations in whole-body RQ, reflective of a metabolically flexible state in which mitochondria switch freely between substrates (fat and sugar) based on nutritional and physiological cues. For the purpose of this discussion, a metabolically sensitive and flexible system is defined as one in which nutrient and energetic signals are rapidly propagated and appropriately interpreted to elicit finely tuned adjustments in fuel partitioning. The physiological importance of metabolic plasticity cannot be understood without first considering the evolutionary pressure for mitochondria to choose fat as a fuel source when systemic glucose reserves are threatened. Whereas lipids provide an abundant, carbon-rich energy source for most tissues, the brain has limited capacity for fat catabolism and therefore relies heavily on a continuous supply of glucose. During periods of food restriction or sustained exercise, protection against hypoglycemia is accomplished by having more versatile tissues (e.g., cardiac, skeletal muscle,

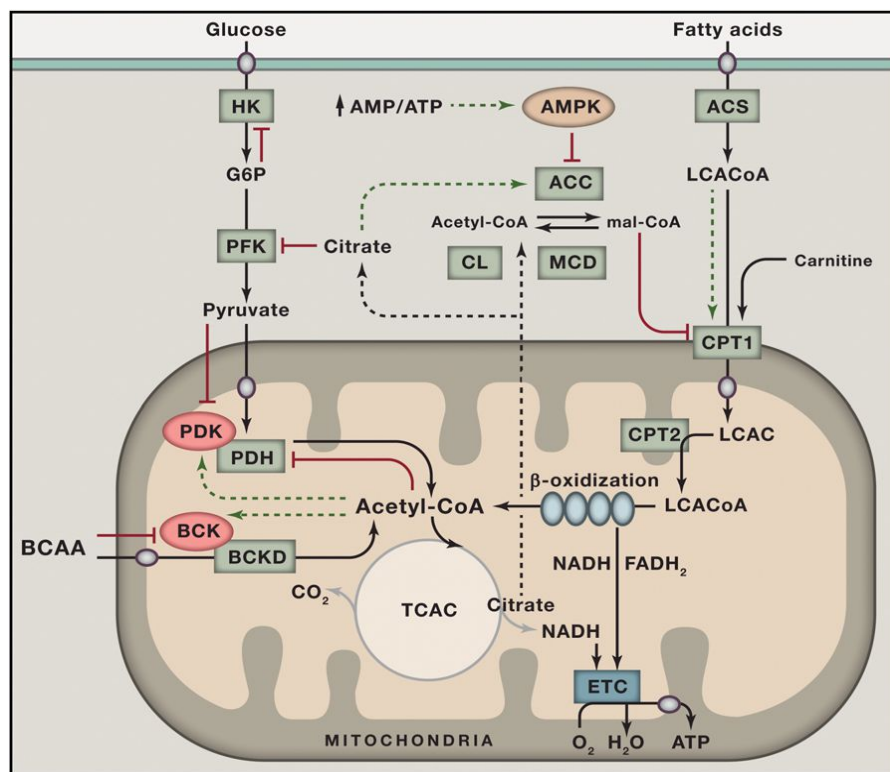


Figure 1. Nutrient Sensing and Signaling Regulate Substrate Selection during Fasting and Feeding

Glucose and fatty acids serve as the primary catabolic substrates that provide acetyl-CoA to the tricarboxylic acid cycle. The pathways of glucose and fat oxidation are reciprocally regulated by several key metabolic intermediates and signals. During fasting, elevated acetyl-CoA derived from high rates of β -oxidation lowers glucose oxidation by allosterically inhibiting PDH and by activating its inhibitory kinase, PDK. Conversely, feeding and glucose surplus restrict fat oxidation by increasing production of malonyl-CoA, which inhibits CPT1. Citrate acts as a signal of plenty that limits glycolytic flux by inhibiting PFK and lowers β -oxidation by giving rise to cytoplasmic acetyl-CoA and malonyl-CoA via CL and ACC, respectively. During periods of energy deficit, an increase in the cellular AMP/ATP ratio activates AMPK, which phosphorylates and inhibits ACC while also activating MCD, thereby relieving malonyl-CoA-mediated inhibition of CPT-1 and promoting fat oxidation. Catabolism of branched-chain amino acids (BCAA) is regulated by BCKD, which is feedback inhibited by acyl-CoA products of the complex due to activation of its inhibitory kinase, BCK. Increased cellular concentrations of pyruvate, BCAA, and fatty acyl-CoAs promote their own catabolism by antagonizing the inhibitory actions of PDK, BCK, and malonyl-CoA, respectively.

ACC, acetyl-CoA carboxylase; ACS, acyl-CoA synthetase; AMPK, 5' AMP-activated kinase; BCAA, branched-chain amino acids; BCKD, branched-chain ketoacid dehydrogenase; BCK, branched-chain ketoacid dehydrogenase kinase; CL, citrate lyase; CPT1, carnitine palmitoyltransferase 1; ETC, electron transport chain; G6P, glucose 6 phosphate; HK, hexokinase; LCAC, long-chain acylcarnitine; LCACoA, long-chain acyl-CoA; MCD, malonyl-CoA decarboxylase; PFK, phosphofructokinase; PDH, pyruvate dehydrogenase; PDK, PDH kinase. Red indicates inhibition; green indicates activation; circles are transporters.

BCKD kinase; CL, citrate lyase; CPT1, carnitine palmitoyltransferase 1; ETC, electron transport chain; G6P, glucose 6 phosphate; HK, hexokinase; LCAC, long-chain acylcarnitine; LCACoA, long-chain acyl-CoA; MCD, malonyl-CoA decarboxylase; PFK, phosphofructokinase; PDH, pyruvate dehydrogenase; PDK, PDH kinase. Red indicates inhibition; green indicates activation; circles are transporters.

liver) convert to a lipid-based respiratory economy. At a systemic level, this switch in metabolic currency is mediated in large part by the counter-regulatory hormones, insulin and glucagon, both of which exert strong influence on adipose tissue lipolysis. Fasting lowers the insulin:glucagon ratio, which stimulates hydrolysis of adipose tissue triacylglycerol and thereby increases delivery of free fatty acids to the periphery. Lipolysis also occurs in tissues such as muscle and liver, further facilitating rapid provision of lipid fuel.

As first recognized by Randle (Randle, 1998), systemic changes in energy supply and demand are also monitored and controlled locally, such that glucose consumption is suppressed when fat oxidation increases. For example, increased supply and oxidation of fatty acids leads to cellular accumulation of acetyl-CoA, NADH, and ATP, which allosterically inhibit pyruvate dehydrogenase (PDH), the mitochondrial enzyme complex that couples glycolysis to glucose oxidation (Figure 1). This same set of allosteric effectors activates a family of PDH kinases that phosphorylate the complex, further inhibiting its catalytic activity (Sugden and Holness, 2006). Randle's glucose-fatty acid hypothesis further proposed that a rise in cellular citrate inhibits phosphofructokinase-1 and that lowering of glycolysis and pyruvate oxidation results in accumulation of glucose-6-phosphate (G6P), leading to allosteric inhibition of hexokinase 2 (HK2) in muscle and heart and diminished glucose uptake. This reciprocal regulation of substrate selection not only preserves

glucose for the brain, but also relieves pyruvate from duties as an oxidative fuel while permitting its use as a gluconeogenic precursor in liver or an anaplerotic substrate that refills the TCAC in muscle and heart.

Another important alternative fuel source during starvation comes as a result of proteolysis and amino acid catabolism, which is regulated, in part, by the mitochondrial branched-chain ketoacid dehydrogenase (BCKD) complex. Like PDH, this complex is allosterically inhibited by NADH and the acyl-CoA esters that arise during branched-chain amino acid catabolism and is covalently inactivated by phosphorylation via BCKD kinase. Also noteworthy, the branched-chain amino-acid-derived α -ketoacid substrates of BCKD inhibit this kinase and thereby promote complex activity when amino acids are present in excess (Shimomura et al., 2001). This regulatory strategy conserves cellular proteins during short periods of fasting and promotes amino acid catabolism in response to a protein-rich diet and during prolonged starvation or exercise when muscle protein breakdown is activated.

By contrast, the nutrient and hormonal milieu elicited by the postprandial state favors glucose uptake, glycolysis, and pyruvate oxidation and a corresponding suppression of fatty acid catabolism. The mechanisms governing the meal-induced switch from fatty acid to glucose oxidation first came to light through the elegant work of McGarry and colleagues (McGarry, 2002), who discovered that feeding increases tissue levels of a

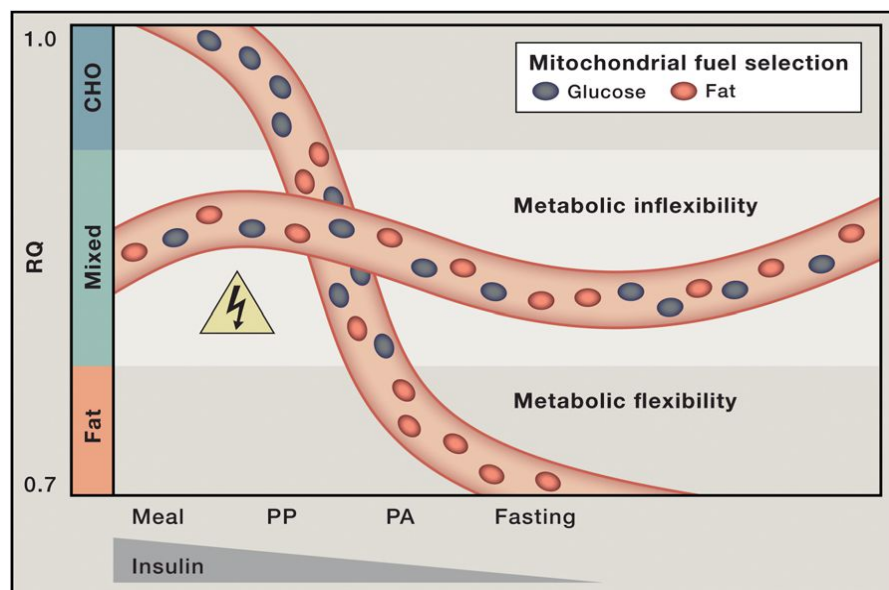


Figure 2. Mitochondrial Indecision Results in Metabolic Inflexibility

In healthy, metabolically flexible states, consumption of a high-carbohydrate (CHO) meal together with a small rise in blood insulin levels elicit a surge in the respiratory quotient (RQ; VCO_2/VO_2), indicative of a robust shift from fatty acid to glucose oxidation. During the postprandial (PP) hours following a meal, mitochondria consume a mixture of fats and carbohydrate. Progression toward the postabsorptive (PA) state and prolonged fasting are accompanied by increased fat oxidation and a corresponding decline in the RQ. Mitochondrial capacity to switch freely between oxidative fuels depending on the nutritional context is lost in obese, metabolically inflexible individuals. Persistent oxidation of a mixture of carbon fuels increases mitochondrial congestion and risk of metabolic hazard.

state. Preferential selection of glucose during the fasted-to-fed transition prevents hyperglycemia, whereas the switch to alternative fuels (lipids and amino

precursor for *de novo* lipogenesis, malonyl-CoA, that also serves as a potent allosteric inhibitor of carnitine palmitoyltransferase-1 (CPT-1) (Figure 1). Positioned on the outer mitochondrial membrane, CPT-1 converts long-chain fatty acyl-CoAs to long-chain acylcarnitines that are able to traverse the inner membrane. Accordingly, CPT-1 acts as the gateway for entry of fatty acids into the mitochondrial matrix. Production of malonyl-CoA, the gatekeeper, increases when glucose is plentiful, thus restricting β -oxidation (McGarry, 2002). Additionally, the glucose-induced rise in pyruvate inhibits the PDKs and thereby favors dephosphorylation and activation of the PDH complex. This step results in refilling of the TCAC and net export of citrate. Upon delivery to the cytoplasm, citrate is cleaved by ATP-citrate lyase to oxaloacetate and acetyl-CoA, the latter of which is then carboxylated and converted to malonyl-CoA by one of two isoforms of acetyl CoA carboxylase (ACC). Citrate also acts as an allosteric activator of ACC in a feedforward manner. In sum, an influx of glucose into the TCAC results in mitochondrial export of carbon metabolites that serve as negative regulators of fat oxidation to facilitate a robust switch in fuel selection. Transitions from a fed state back to a state of energy deficit (e.g., fasting or exercise) relieve the inhibition fat oxidation through activation of the energy stress sensor, 5' AMP-activated kinase (AMPK), which phosphorylates and inactivates ACC, thereby lowering malonyl-CoA levels and increasing CPT-1 activity.

Collectively, the mechanisms originally described by Randle and McGarry represent key components of a sophisticated metabolic network that monitors and responds to the local and systemic nutrient environments to maintain glucose homeostasis. Transitions between fasting and feeding trigger a wave of metabolite signals that guide mitochondrial fuel choice and regulate carbon trafficking (Figure 1). Metabolic intermediates arising from fat catabolism act as negative regulators of glucose oxidation and vice versa. Crosstalk and cooperation between competing substrates enable mitochondria to choose the energy source that is most appropriate for a particular physiological

acids) during periods of food deprivation defends against hypoglycemia and ensures organism survival.

Metabolic Congestion Results in Mitochondrial Indecision

It is important to consider that human physiology evolved to cope with dramatic fluctuations in energy supply and demand during periods of feast and famine or hunting/gathering. Thus, episodes of refueling were typically preceded by a sustained period of energy deficit. By contrast, physiology in the modern era is characterized by a steady influx of competing fuels. A large body of evidence suggests that overnutrition and unabated substrate competition lead to a state of metabolic insensitivity and inflexibility, characterized by distorted nutrient sensing, blunted substrate switching, and impaired energy homeostasis.

The concept of metabolic inflexibility was first introduced by Kelley and colleagues, who monitored gas exchange across the leg to examine substrate switching in healthy compared to obese and diabetic subjects (Figure 2) (Kelley et al., 1999). Lean, healthy subjects shifted from a low RQ in the fasted state to a high RQ during a hyperinsulinemic-euglycemic clamp, an infusion procedure that mimics the fed state by increasing plasma insulin concentrations while holding blood glucose constant at basal levels. When this same test was applied to obese and type 2 diabetic subjects, the transition to the "fed" state was accompanied by only a marginal change in RQ. Therefore, the insulin-resistant individuals continued to oxidize a fixed mixture of fats and carbohydrates regardless of the nutritional context. Additionally, emerging evidence suggests that increased amino acid supply and catabolism (which elicits an intermediate RQ) might also contribute to obesity-related perturbations in fuel use (Newgard, 2012). Thus, in the context of chronic overfeeding, competition between substrates escalates, cooperation is lost, and the mitochondria are left in a state of indecision characterized by persistent oxidation of all three major fuels. This phenomenon of blunted fuel switching has now been described

in a variety of clinical settings, including obesity (Prior et al., 2014), diabetes (Ukropcova et al., 2007), heart disease (Turer et al., 2010), nonalcoholic steatohepatitis (Lee et al., 2014), polycystic ovarian syndrome (Di Sarra et al., 2013), and physical inactivity (Bergouignan et al., 2013). Whereas most of these studies evaluated substrate use by skeletal muscle, heart, and/or liver, emerging evidence shows that similar perturbations in fuel switching and nutrient responsiveness manifest in other organs and cell types, including adipose tissue (Sparks et al., 2009), macrophages (Asterholm et al., 2012a), and monocytes (Liu et al., 2012).

Metabolic inflexibility can also be induced experimentally in rodents with high-fat feeding (Koves et al., 2008; Newgard, 2012) or as a consequence of one of numerous and disparate genetic manipulations used to mimic various metabolic disease states (Asterholm et al., 2012b; Koves et al., 2008; Vadvalkar et al., 2013). In many cases, mitochondrial capacity to switch between fuels is coupled to changes in cellular and/or tissue functions, such as insulin action, glucose disposal, lipolysis, lipid storage, cardiac contractility, immune function, and inflammatory response (Asterholm et al., 2012a, 2012b; Koves et al., 2008; Liu et al., 2012; Sparks et al., 2009; Vadvalkar et al., 2013). Moreover, sluggish responses to a change in nutrient load appear to extend beyond substrate selection and are observed at the level of gene and protein expression. Thus, robust induction or suppression of a wide range of transcripts, which typically occurs during the fasting-to-fed transition, is attenuated in models of disease (Gao et al., 2014; Jans et al., 2011). This consequence is not surprising considering that many of the aforementioned signaling events and metabolic intermediates that mediate the glucose-fatty acid cycle also influence gene transcription via direct or indirect mechanisms. Although this mode of regulation is unlikely to influence immediate responses to a meal and/or insulin stimulation, nutrient-induced modulation of mRNA and protein abundance probably reflects cellular anticipation of prolonged stress and/or priming of the network for the next meal. This type of hormetic adaptation builds regulatory reserve and enhances the capacity of the network to monitor and cope with future metabolic insults. Accordingly, perturbations in meal-induced transcriptional/translational reprogramming might compromise nutrient sensitivity in response to chronic metabolic pressures (Gao et al., 2014).

Mitochondrial Overload Leads to Metabolic Gridlock

Research to delineate the molecular origins of metabolic inflexibility has focused largely on the glucose-fatty acid cycle and/or aberrant production of malonyl-CoA. Although these mechanisms certainly weigh heavily on fuel selection, the finding that inflexibility is associated with or can be provoked by a broad range of clinical and experimental circumstances suggests that the molecular basis of this condition reaches beyond dysregulation of PDH and/or CPT-1. It is important to consider that: (1) fuel selection occurs at the level of the mitochondrion, (2) carbon substrates are the source of the electrons that feed the ETC, and (3) oxidative phosphorylation satisfies 70%–90% of ATP demand in most cells. Accordingly, this organelle is ideally positioned to monitor and transmit energy and nutrient status

throughout the metabolic network (Anderson et al., 2009; Mailoux et al., 2013; Muoio and Neuffer, 2012; Newman et al., 2012). This mito-centric model of nutrient sensing and partitioning suggests that cellular energy charge and shifts in flux control are integrated and executed as a function of mitochondrial carbon load.

To conceptualize a network model of metabolic inflexibility, the spatiotemporal features of carbon flux through the metabolic interstates of an organism transitioning from fasting to feeding can be viewed in a manner analogous to the onset of rush hour traffic. When volume is light, free-flowing traffic remains highly responsive to internal cues based on operator decisions and mechanical function, as well as external inputs from traffic signs and signals. Under these circumstances, distance between vehicles remains generous, buffering capacity is robust, and abrupt fluctuations in traffic density and speed can occur with minimal risk of collision. As volume expands, bottlenecks at highly traveled intersections impede flow, the buffering capacity of the network diminishes, and the probability of random collision mounts. An incident at a critical node can lead to systemic paralysis, such that traffic flow remains unresponsive even when a major signal changes from red to green. The roadways reach a state of gridlock, and the time and energy required to restore normal flow depends on the extent of the impasse and the severity of the damage.

It is conceivable that a similar state of gridlock develops when organisms eat voraciously and often (Figure 3). During and immediately after each meal, carbon traffic becomes more congested and competition between substrates intensifies. Moreover, as intracellular triacylglycerol and glycogen depots reach capacity, continuous turnover of these large reservoirs imposes additional nutrient pressure on the network due to mass action. Eventually, carbon flux is perturbed, not only due to the high volume of traffic, but also as a result of distorted and conflicting signals that provoke a situation of metabolic “road rage” between the products of CPT-1, PDH, and BCKD. As tensions escalate and the three substrates battle for the right of way, mitochondria are confronted with an ever-increasing nutrient burden.

Based on first-order kinetics, the overfed mitochondria continue to degrade incoming carbon substrates. Although rates of fuel catabolism might be low, each molecule of acetyl-CoA produced by the processes of glucose, fat, and amino acid oxidation is accompanied by the generation and delivery of reducing equivalents to the ETC. Because OXPHOS is a demand-driven process regulated by ADP availability, an increase in electron delivery does not necessitate a proportional increase in ATP production. When electron supply to the Q cycle exceeds demand for ATP synthesis, mitochondrial membrane potential rises and proton pumping at complexes I, III, and IV is met with increasing back pressure. Under these circumstances, the main escape route for incoming electrons occurs via the reduction of molecular oxygen and generation of the superoxide anion, followed by its rapid conversion to hydrogen peroxide (H₂O₂) by superoxide dismutase. Recent studies also identify PDH, BCKD, and α -ketoglutarate dehydrogenase as important sites of ROS production (Fisher-Wellman et al., 2013). Consequently, as the NADH/NAD⁽⁺⁾ redox pair shifts to a more reduced state, the microenvironments surrounding the ETC and the 2-oxoacid

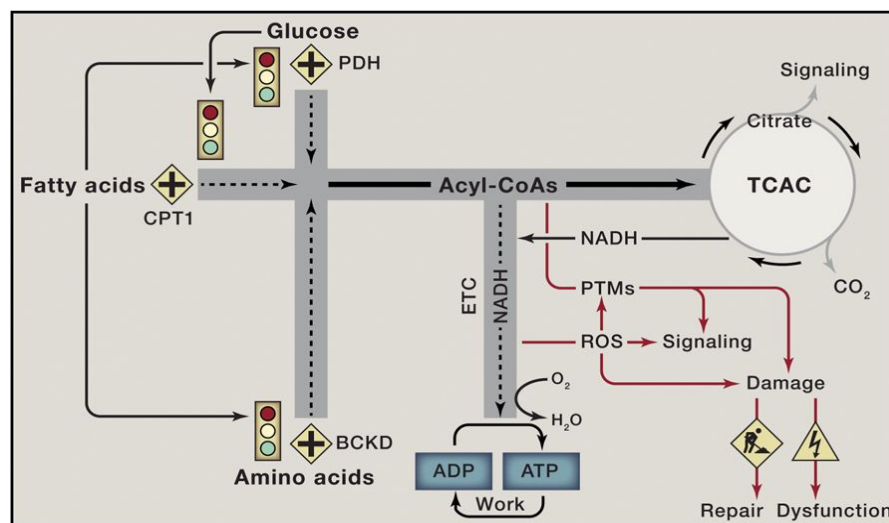


Figure 3. Nutrient Overload Leads to Mitochondrial Gridlock and Cellular Dysfunction

Glucose, fatty acids, and branched-chain amino acids are degraded to acyl-CoA intermediates that fuel the mitochondrial tricarboxylic acid cycle (TCAC). All catabolic roadways in the mitochondria, including the TCAC, lead to the production of reducing equivalents (NADH) that feed the electron transport chain (ETC), thereby permitting ATP regeneration to support cellular work and energy expenditure. In healthy states, carbon traffic at the crossroads marked by PDH, CPT1, and BCKD is coordinately and reciprocally regulated by an intricate network of metabolic signals that match energy supply to ATP demand. High rates of fat oxidation inhibit glucose and BCAA catabolism and vice versa, thereby preventing mitochondrial congestion when ATP consumption is low. Chronic overnutrition causes metabolic confusion and signal failure, resulting in unabated influx of surplus fuel and an ensuing traffic jam at several critical bottlenecks where the roadways converge. As mitochondrial traffic reaches a state of gridlock, membrane potential rises and accumulating electrons and acyl-CoAs are diverted toward ROS

generation and PTMs such as glutathionylation and lysine acetylation, which further disrupts nutrient sensing and signaling. If these road hazards are not sufficiently managed by the mitochondrial buffering and repair systems, mounting irreversible damage to cellular macromolecules leads to organ dysfunction.

dehydrogenase complexes becomes more conducive to H_2O_2 production and emission (Anderson et al., 2009).

Hydrogen peroxide and other reactive oxygen species are now well recognized as bona fide signaling molecules that modulate reversible oxidation/reduction of sulfur atoms within critical cysteine residues of numerous proteins. The interconversion of these so-called “sulfur switches” from protein thiols (reduced form) to their corresponding disulfides (oxidized form) affects the activities and/or functions of an expansive network of redox-sensitive metabolic enzymes and signaling proteins (Brandes et al., 2009). This link implies that perturbations in the frequency, amplitude, and/or duration of the mitochondrial H_2O_2 pulse could have far-reaching effects on redox circuitry, nutrient flux, and energy homeostasis (Mailloux et al., 2013). Additionally, as oxidative stress builds, H_2O_2 and other ROS are more likely to collide with and damage cellular constituents via irreversible reactions such as protein carbonylation and lipid peroxidation (Frohnert and Bernlohr, 2013), further compromising the integrity and plasticity of the network.

Meanwhile, as delivery of carbons persists and the amount of excess fuel mounts, the NADH/NAD⁽⁺⁾ ratio within the mitochondrial lumen increases and redox inhibition of several TCAC enzymes limits flux through this major metabolic beltway. The ensuing mismatch between the early steps of carbon degradation and TCAC flux can lead to intramitochondrial accumulation of acetyl-CoA and other acyl-CoAs at several crucial bottlenecks, particularly at sites where catabolism of the three fuels converge (Figure 3) (Koves et al., 2008; Newgard, 2012). Also notable is that citrate synthase, the main entry point into the TCAC, does not bind acetyl-CoA without first binding to its other substrate, oxaloacetate. Additionally, amino-acid-derived succinyl-CoA acts as a potent inhibitor of citrate synthase. As such, a limitation at the level oxaloacetate and/or accumulation of succinyl-CoA adds further pressure on the expanding mitochondrial pool of acetyl-CoA.

Acetyl-CoA and other reactive thioesters not only act as allosteric regulators of mitochondrial enzymes, but also serve as acyl donors for protein modifications (PTMs) such as lysine acetylation, succinylation, and palmitoylation via either enzymatic or nonenzymatic mechanisms (Choudhary et al., 2014; Wagner and Payne, 2013). Lysine acetylation is a reversible PTM in which a two carbon acetyl group is covalently attached to the ϵ -amino group of a lysine residue. Mass-spectrometry-based acetyl-proteomic analyses have led to the estimate that more than one-third of the mitochondrial proteome is acetylated on at least one lysine residue, apparently affecting nearly every major pathway of intermediary metabolism. Moreover, global acetylation in liver and muscle increases in response to diet-induced obesity in mice. Although the functional relevance of these PTMs is largely undefined, a growing number of mitochondrial enzymes have been shown to be negatively regulated by lysine acetylation (Choudhary et al., 2014; Gao et al., 2014; Hirschey et al., 2010; Jing et al., 2011; Jing et al., 2013). Yet unclear is whether these PTMs represent bona fide signaling events and/or a form of protein damage that comes as a cost of traveling the mitochondrial roadways during rush hour traffic. Either way, a substantive shift in the mitochondrial acylome might perturb carbon and electron flow, leading to sluggish metabolic responses to nutritional and/or hormonal stimuli.

Metabolic Countermeasures and Mitochondrial Damage Control

Considering that most routes connecting food consumption to ATP production travel through the mitochondria and because reactive lysine and cysteine residues of mitochondrial proteins are more vulnerable to nucleophilic attack due to the alkaline environment of the matrix (Ghanta et al., 2013; Mailloux et al., 2013; Wagner and Payne, 2013), it is not surprising that this organelle has developed various countermeasures to buffer excess traffic and repair the damage caused by inadvertent

molecular collisions. For example, the redox circuits modulated by H_2O_2 are buffered by the interdependent glutathione- and thioredoxin-reducing systems (Mailloux et al., 2013). Both use the reducing power of NADPH to mitigate oxidative stress and to modulate reversible oxidation/reduction of protein thiols/disulfides. Additionally, oxidized glutathione molecules (GSSG) can form disulfide linkages with reactive cysteine residues within the redox-sensitive proteome (Mailloux et al., 2013). This interaction produces a PTM known as S-glutathionylation (PSSG), which is thought to protect proteins from permanent oxidative damage. S-glutathionylation of mitochondrial proteins under normal conditions is cysteine thiol specific, reversible, sensitive to changes in redox environment, and enzyme driven (Mailloux et al., 2013). Thus, the S-glutathionylation cycle, catalyzed by a family of protein redoxins, acts as a redox rheostat that modulates the function of a large number of protein targets, including mitochondrial and antioxidant enzymes and several protein kinases and phosphatases (Mailloux et al., 2013). Relevant to this area of study, numerous reports show that circulating and/or tissue levels of GSH decline in the context of obesity and metabolic disease (Anderson et al., 2009), suggesting diminished redox buffering capacity.

Another buffering system that plays a key role in energy homeostasis utilizes a family of mitochondrial carnitine acyltransferase (CAT) enzymes that catalyze the exchange of acyl groups between CoA and L-carnitine (Ramsay and Zammit, 2004). Unlike their acyl-CoA precursors, acylcarnitine esters are readily transported across cellular membranes. Accordingly, this system permits shuttling of carbon fuels between compartments and thereby acts to offset nutrient-induced expansion of the mitochondrial pools of acetyl- and other reactive acyl-CoA moieties. With the advent and broadening applications of metabolomics technologies, acylcarnitine metabolites have emerged as strong biomarkers of nutrient stress, mitochondrial dysfunction, and metabolic disease (Koves et al., 2008; Muoio and Neuffer, 2012; Newgard, 2012; Noland et al., 2009). The most abundant acylcarnitine species, acetylcarnitine, is synthesized by carnitine acetyltransferase (CrAT), a member of the CAT family that localizes to the mitochondrial matrix and strongly prefers acetyl- and other short-chain acyl-CoA end products of fatty acid, glucose, and amino acid catabolism (Muoio et al., 2012). Notably, muscle-specific ablation of *crat* in mice elevates tissue levels of acetyl-CoA, lowers PDH activity, increases acetylation of several mitochondrial proteins, and diminishes metabolic flexibility (Muoio et al., 2012). Conversely, dietary L-carnitine supplementation promotes glucose tolerance and enhances metabolic flexibility in concert with increased circulating levels of the main CrAT product, acetylcarnitine (Muoio et al., 2012; Noland et al., 2009). Likewise, muscle CrAT activity correlates positively with insulin sensitivity in rodents and humans (Lindeboom et al., 2014). These studies establish an important link between acylcarnitine efflux, mitochondrial acyl-CoA balance, and fuel selection and strongly imply that mitochondrial carbon load directly impacts glucose tolerance and insulin action.

Whereas the carnitine system relieves substrate push on mitochondrial protein acylation, the sirtuin family of NAD^+ -dependent deacylases performs damage control by removing acyl groups from lysine residues. Importantly, however, the sirtuins do not

act on all acyl-lysine residues; thus, it appears that only a specific subset of these PTMs is reversible. SIRT3, SIRT4, and SIRT5 are found in the mitochondrial matrix. SIRT3 is the main mitochondrial deacetylase and the best characterized of the mitochondrial sirtuins (Newman et al., 2012). SIRT5 has robust desuccinylase, demalonylase, and deglutarylase activities, whereas most sirtuins (SIRT1–6) appear to also possess long-chain deacylation activity (Feldman et al., 2013). Interestingly, Sirt3 knockout mice have increased acetylation of several metabolic enzymes, including PDH and superoxide dismutase, which is accompanied by elevated acylcarnitines and increased mitochondrial ROS production (Hirschey et al., 2010; Jing et al., 2011; Jing et al., 2013). These animals are less metabolically flexible than their wild-type counterparts and also exhibit multiple features of the metabolic syndrome as they age. Because the lysine deacylation reactions consume NAD^+ , which is limiting for sirtuin activity, obesity-induced lowering of NAD^+ is thought to constrain the sirtuin system (Peek et al., 2013). Thus, the combination of increased acetyl donors and reduced deacylating capacity might underlie the elevation in mitochondrial protein acetylation observed in response to chronic nutrient overload. These findings have fueled strong interest in the use of nicotinamide riboside and other NAD^+ precursors as an antidiabetic strategy to boost sirtuin activity and restore metabolic function (Cantó et al., 2012).

Insulin Resistance Viewed as a Case of Mixed Signals

Aberrant transitions between fuel types at both the cellular and systemic levels link to phenotypes associated with metabolic comorbidities, including insulin resistance. An important question is whether insulin resistance causes metabolic inflexibility or vice versa. One view suggests that blunted glucose oxidation in response to a meal simply reflects a consequence of impaired insulin signaling. On the other hand, studies in rodents show that metabolic inflexibility occurs early in the course of glucose intolerance, and obesity-induced perturbations in substrate switching are evident in isolated mitochondria and tissue homogenates (Muoio et al., 2012; Noland et al., 2009). These findings suggest that derangements in fuel selection are at least partly independent of and might actually precede and contribute to insulin resistance.

It is important to reiterate that the models developed by Randle and McGarry center on the concepts of reciprocation, cooperation, and communication between substrates. In a healthy state, crosstalk between metabolic pathways is mediated by robust, concise, and decisive changes in cellular levels of metabolites such as fatty acids, pyruvate, citrate, and malonyl-CoA, which in turn regulate incoming mitochondrial traffic (Figure 1). By contrast, chronic overnutrition leads to a state of metabolic confusion, wherein excessive carbon supply and heightened substrate competition give rise to a set of muted and/or conflicting signals. As a result, the gateways that control mitochondrial traffic are never fully open or shut, and the continuous influx of carbon fuel from multiple sources and directions interferes with efficient substrate switching. Thus, metabolic inflexibility can be viewed as both a cause and indicator of mitochondrial congestion, which in turn influences insulin action. Transitions between highly insulin-sensitive and more

insulin-resistant states appear to operate on a continuum as a function of carbon and electron supply relative to ATP demand (Figure 3), such that a shift in the positive direction triggers a set of mitochondrial-derived stress signals that oppose cellular glucose uptake as an attempt to limit traffic volume.

From an evolutionary perspective, it seems probable that more primitive and firmly hardwired intracellular nutrient-sensing networks of the cell override hormonal stimuli when the signals disagree. One example of this hierarchy comes from studies of exercise and muscle contraction. Although the systemic nutrient and hormonal environment of exercise mimics a starvation state on many levels, glucose flux into muscle during contraction proceeds unimpeded. The energy charge of working muscles stimulates translocation of the resident glucose transporter, Glut4, from an intracellular compartment to the cell surface, thereby permitting glucose uptake. Whereas Glut4 functions as the classical insulin-responsive glucose transporter, its translocation during exercise occurs via an insulin-independent pathway involving short-acting signals such as an elevated AMP/ATP ratio and consequent activation of AMPK (Jessen and Goodyear, 2005). Similarly, numerous studies have shown that a single bout of vigorous exercise enhances insulin action for up to 24 hr in muscles of obese rodents or humans (Thyfault, 2008). These observations underscore two important points. First, obesity does not cause insulin resistance. Instead, increased adiposity is merely a symptom of chronic (positive) energy imbalance, presumably the true culprit. Second, even brief episodes of increased energy expenditure and accelerated carbon combustion can reset the energy charge of the muscle and enhance insulin action. The precise molecular mechanisms underlying exercise-mediated enhancement of insulin responsiveness have evaded scientists for decades. Interestingly, however, recent studies have established a strong positive association between exercise training and metabolic flexibility (Bergouignan et al., 2013; Koves et al., 2013), suggesting that physical activity enhances mitochondrial traffic control and that persistent substrate competition is a key component of insulin resistance.

Also noteworthy is that a number of reports have found that glucose intolerance precedes overt defects in the canonical insulin signaling cascade leading to Akt/PKB (protein kinase B) phosphorylation (Muoio and Neufer, 2012; Muoio et al., 2012). Thus, intracellular cues dissuading glucose uptake appear to take precedence over Akt activation. This state might be partly attributable to lipid-induced accumulation of G6P and/or flux limitations at HK2, as proposed by Randle and others (Randle, 1998; Wasserman et al., 2011), but likely involves additional feedback signals that further antagonize glucose disposal (Muoio and Neufer, 2012). In this scenario, the pancreatic β cells should compensate for diminished rates of peripheral glucose disposal by releasing more insulin, resulting in hyperinsulinemia. Over time, this compensatory response could lead to insulin-mediated insulin resistance whereby insulin itself is responsible for desensitizing the insulin receptor and its proximal targets (Copps and White, 2012).

Additionally, heightened and persistent competition between oxidative substrates has the potential to impact several other extra-mitochondrial signaling molecules known or presumed to influence insulin action. For instance, diminished flux through

PDH could result in re-routing of glycolytic intermediates toward gluconeogenesis in liver or the generation of lipid-signaling molecules such as diacylglycerol and ceramide, all of which have been strongly implicated in insulin resistance (Muoio and Newgard, 2008). In general, excessive production of these lipid species has been attributed to fatty acid toxicity. Notably, however, *de novo* synthesis of both molecules requires glucose-derived carbon intermediates—namely, glycerol-3-phosphate and serine, respectively. In other words, “it takes two to tango”; therefore, cellular synthesis of diacylglycerols and ceramides diminishes when one of the two precursor fuels is fully committed to the mitochondria for the purpose of ATP production.

As noted earlier, heightened substrate competition promotes ROS production. Mitochondrial-derived H_2O_2 can oxidize critical cysteine residues within the catalytic sites of several protein phosphatase enzymes that modulate insulin signaling in both directions, including the phosphoinositide phosphatase, PTEN, and several dual specificity protein tyrosine phosphatases. Thus, perturbations in the pattern of H_2O_2 generation could lead to over activation and/or desensitization of the insulin signaling network. Along with alterations in protein function, major shifts in the redox proteome and/or other PTMs, including carbonylation, glutathionylation, and lysine acylation, might disrupt protein turnover and/or folding (Frohnert and Bernlohr, 2013; Gao et al., 2014; Mailloux et al., 2013). Perturbations in cellular proteostasis trigger the unfolded protein response (UPR) and endoplasmic reticulum stress, another candidate mediator of insulin resistance (Muoio and Newgard, 2008). Moreover, emerging evidence suggests that hyperacylation of the mitochondrial proteome influences mitophagy, mitochondrial dynamics, and the mitochondrial UPR (Gao et al., 2014; Papa and Germain, 2014). Lastly, several of the foregoing PTMs have also been implicated in the regulation/dysregulation of various mitochondrial membrane organic acid carrier proteins and other metabolite transporters, including the citrate carrier, the adenine nucleotide translocase, and the carnitine acylcarnitine translocase. Thus, these modifications have the potential to obstruct crosstalk between mitochondria and other cellular compartments, which could dampen meal-induced shifts in nutrient partitioning, glucose flux, and retrograde signaling.

Strategies for Decongestion

The idea that mitochondrial carbon overload and metabolic inflexibility lie at the core of insulin resistance implies that maneuvers to prevent oxidative catabolism of at least one of the three major fuel sources might alleviate substrate competition and restore glucose control. This strategy has, in fact, proven beneficial in mouse models with genetically engineered inhibition of CPT-1 or other proximal steps in fatty acid oxidation (Koves et al., 2008; Li et al., 2014; Muoio and Neufer, 2012). Although fewer studies have targeted glucose oxidation, there is one report showing that cardiac-specific overexpression of PDK4 caused marked suppression of heart PDH activity and glucose oxidation but without the expected deleterious functional and metabolic consequences in response to an ischemia-reperfusion challenge (Chambers et al., 2011). Numerous provocative studies in animals and humans show that very low-carbohydrate (ketogenic) diets produce favorable metabolic outcomes when

compared to a traditional mixed macronutrient diet (Hu et al., 2012). Whereas these findings have sparked intense controversy over dietary recommendations for disease prevention, perhaps the salient observation is that a dietary regimen that effectively limits mitochondrial substrate competition produces marked improvements in metabolic control. Alternatively, several lines of evidence suggest that boosting or restoring the buffering capacities of the systems that mitigate mitochondrial carbon stress, including antioxidant defense, glutathione regeneration, carnitine-mediated acyl group buffering, and sirtuin-mediated protein deacylation, also improves whole-body energy homeostasis and glucose control in association with enhanced metabolic flexibility (Anderson et al., 2009; Cantó et al., 2012; Muoio et al., 2012; Noland et al., 2009).

Although blocking mitochondrial influx of carbons and/or building buffering capacity might alleviate the metabolic load on this particular organelle, these strategies do not address the fundamental problem of energy imbalance and the broader implications of nutrient surplus. Thus, approaches targeting the cause rather than the symptoms of nutrient overload should yield more desirable outcomes. Moving beyond the standard recommendation to “eat less and move more,” recent studies employing unconventional behavior modification interventions have produced intriguing results, including intermittent fasting regimens in which rodents or human subjects fast for extended periods followed by ad libitum eating (Azevedo et al., 2013; Trepanowski et al., 2011), as well as “exercise snacking” regimens in which individuals exercise vigorously for brief (~5 min) periods before consuming a standard-size meal (Francois et al., 2014). Interestingly, these routines were found to improve glucose homeostasis more than isoenergetic regimens comprised of traditional feeding and exercise patterns. Likewise, interruption of prolonged sedentary behavior with repeated bouts of low-grade activities of daily living (e.g., post-meal strolling) has been shown to improve postprandial glucose handling in patients with type 2 diabetes (van Dijk et al., 2013). These observations are consistent with the premise that periodic episodes of accelerated carbon combustion, in a manner resembling the lifestyle of our hunter/gatherer ancestors, facilitate rapid and efficient nutrient partitioning.

Relevant to this discussion is the prominent connection between caloric restriction, lifespan, and healthspan. In model organisms, lifelong caloric restriction increases longevity and delays age-related metabolic decline. Similar outcomes have been reported in nonhuman primates, and epidemiological studies have revealed a strong association between energy balance and healthspan in humans (Trepanowski et al., 2011). The caloric restriction literature implies that ad libitum feeding is inevitably damaging to metabolic health and that surplus fuel, beyond that required for optimal cellular function, accelerates biological decay. Energy balance is typically considered at a macro level over periods of days, weeks, or months, with body weight and adiposity targeted as the primary readouts. Alternatively, tracking of nutrient balance at the cellular level on a minute-by-minute or hourly basis might represent the more biologically relevant scale. In theory, each meal tips carbon balance of each individual cell in a positive direction. Whereas the incoming nutrients are ultimately destined for either biosynthetic

or catabolic fates, slow or misregulated decisions on how to appropriate the extra energy (i.e., metabolic inflexibility) can result in collateral damage. This paradigm aligns with emerging evidence that larger, less frequent meals are more detrimental to nutrient sensing, retrograde signaling, and metabolic control than smaller, more frequent meals (Fuse et al., 2012; Trepanowski et al., 2011).

Concluding Remarks

In summary, this discussion speaks to the longstanding question of whether fuel selection matters to metabolic health. Historically, this area of study has been divided into opposing viewpoints that argue for or against the idea that it is healthier to burn fat than glucose. In recent years, these two camps have found common ground in the concept of metabolic flexibility, which posits that cells function optimally when they retain their capacity to switch freely between oxidative substrates in response to nutritional and physiological cues. The foregoing network model of metabolic flexibility further suggests that, during conditions of inactivity and low ATP demand, mitochondria function best when acetyl-CoA is produced from one fuel at a time. Robust and decisive shifts in substrate choice are predicted to limit mitochondrial congestion and damaging molecular collisions while also producing strong and clearly interpretable metabolic signals that guide efficient nutrient partitioning to maintain energy homeostasis. This model provides an explanatory context for viewing the severe costs of excessive food consumption and the benefits of habitual physical activity and lifelong caloric restriction.

ACKNOWLEDGMENTS

The author is supported by grants from the United States Public Health Service: R01DK089312, P01DK058398, and R01HL101189.

REFERENCES

- Anderson, E.J., Lustig, M.E., Boyle, K.E., Woodlief, T.L., Kane, D.A., Lin, C.T., Price, J.W., 3rd, Kang, L., Rabinovitch, P.S., Szeto, H.H., et al. (2009). Mitochondrial H₂O₂ emission and cellular redox state link excess fat intake to insulin resistance in both rodents and humans. *J. Clin. Invest.* **119**, 573–581.
- Asterholm, I.W., McDonald, J., Blanchard, P.G., Sinha, M., Xiao, Q., Mistry, J., Rutkowski, J.M., Deshaies, Y., Brekken, R.A., and Scherer, P.E. (2012a). Lack of “immunological fitness” during fasting in metabolically challenged animals. *J. Lipid Res.* **53**, 1254–1267.
- Asterholm, I.W., Mundy, D.I., Weng, J., Anderson, R.G., and Scherer, P.E. (2012b). Altered mitochondrial function and metabolic inflexibility associated with loss of caveolin-1. *Cell Metab.* **15**, 171–185.
- Azevedo, F.R., Ikeoka, D., and Caramelli, B. (2013). Effects of intermittent fasting on metabolism in men. *Rev. Assoc. Med. Bras.* **59**, 167–173.
- Bergougnan, A., Antoun, E., Momken, I., Schoeller, D.A., Gauquelin-Koch, G., Simon, C., and Blanc, S. (2013). Effect of contrasted levels of habitual physical activity on metabolic flexibility. *J. Appl. Physiol.* (1985) **114**, 371–379.
- Brandes, N., Schmitt, S., and Jakob, U. (2009). Thiol-based redox switches in eukaryotic proteins. *Antioxid. Redox Signal.* **11**, 997–1014.
- Cantó, C., Houtkooper, R.H., Pirinen, E., Youn, D.Y., Oosterveer, M.H., Cen, Y., Fernandez-Marcos, P.J., Yamamoto, H., Andreux, P.A., Cettour-Rose, P., et al. (2012). The NAD(+) precursor nicotinamide riboside enhances oxidative metabolism and protects against high-fat diet-induced obesity. *Cell Metab.* **15**, 838–847.

- Chambers, K.T., Leone, T.C., Sambandam, N., Kovacs, A., Wagg, C.S., Lopaschuk, G.D., Finck, B.N., and Kelly, D.P. (2011). Chronic inhibition of pyruvate dehydrogenase in heart triggers an adaptive metabolic response. *J. Biol. Chem.* 286, 11155–11162.
- Choudhary, C., Weinert, B.T., Nishida, Y., Verdin, E., and Mann, M. (2014). The growing landscape of lysine acetylation links metabolism and cell signalling. *Nat. Rev. Mol. Cell Biol.* 15, 536–550.
- Copps, K.D., and White, M.F. (2012). Regulation of insulin sensitivity by serine/threonine phosphorylation of insulin receptor substrate proteins IRS1 and IRS2. *Diabetologia* 55, 2565–2582.
- Di Sarra, D., Tosi, F., Bonin, C., Fiers, T., Kaufman, J.M., Signori, C., Zambotti, F., Dall'Alda, M., Caruso, B., Zanolini, M.E., et al. (2013). Metabolic inflexibility is a feature of women with polycystic ovary syndrome and is associated with both insulin resistance and hyperandrogenism. *J. Clin. Endocrinol. Metab.* 98, 2581–2588.
- Feldman, J.L., Baeza, J., and Denu, J.M. (2013). Activation of the protein deacetylase SIRT6 by long-chain fatty acids and widespread deacylation by mammalian sirtuins. *J. Biol. Chem.* 288, 31350–31356.
- Fisher-Wellman, K.H., Gilliam, L.A., Lin, C.T., Cathey, B.L., Lark, D.S., and Neuffer, P.D. (2013). Mitochondrial glutathione depletion reveals a novel role for the pyruvate dehydrogenase complex as a key H₂O₂-emitting source under conditions of nutrient overload. *Free Radic. Biol. Med.* 65, 1201–1208.
- Francois, M.E., Baldi, J.C., Manning, P.J., Lucas, S.J., Hawley, J.A., Williams, M.J., and Cotter, J.D. (2014). 'Exercise snacks' before meals: a novel strategy to improve glycaemic control in individuals with insulin resistance. *Diabetologia* 57, 1437–1445.
- Frohnert, B.I., and Bernlohr, D.A. (2013). Protein carbonylation, mitochondrial dysfunction, and insulin resistance. *Adv. Nutr.* 4, 157–163.
- Fuse, Y., Hirao, A., Kuroda, H., Otsuka, M., Tahara, Y., and Shibata, S. (2012). Differential roles of breakfast only (one meal per day) and a bigger breakfast with a small dinner (two meals per day) in mice fed a high-fat diet with regard to induced obesity and lipid metabolism. *J. Circadian Rhythms* 10, 4.
- Gao, A.W., Canto, C., and Houtkooper, R.H. (2014). Mitochondrial response to nutrient availability and its role in metabolic disease. *EMBO Mol. Med.* 6, 580–589.
- Ghanta, S., Grossmann, R.E., and Brenner, C. (2013). Mitochondrial protein acetylation as a cell-intrinsic, evolutionary driver of fat storage: chemical and metabolic logic of acetyl-lysine modifications. *Crit. Rev. Biochem. Mol. Biol.* 48, 561–574.
- Hirschey, M.D., Shimazu, T., Goetzman, E., Jing, E., Schwer, B., Lombard, D.B., Grueter, C.A., Harris, C., Biddinger, S., Ilkayeva, O.R., et al. (2010). SIRT3 regulates mitochondrial fatty-acid oxidation by reversible enzyme deacetylation. *Nature* 464, 121–125.
- Hu, T., Mills, K.T., Yao, L., Demanelis, K., Eloustaz, M., Yancy, W.S., Jr., Kelly, T.N., He, J., and Bazzano, L.A. (2012). Effects of low-carbohydrate diets versus low-fat diets on metabolic risk factors: a meta-analysis of randomized controlled clinical trials. *Am. J. Epidemiol.* 176 (Suppl 7), S44–S54.
- Jans, A., Sparks, L.M., van Hees, A.M., Gjelstad, I.M., Tierney, A.C., Risérus, U., Dreven, C.A., Roche, H.M., Schrauwen, P., and Blaak, E.E. (2011). Transcriptional metabolic inflexibility in skeletal muscle among individuals with increasing insulin resistance. *Obesity (Silver Spring)* 19, 2158–2166.
- Jessen, N., and Goodyear, L.J. (2005). Contraction signaling to glucose transport in skeletal muscle. *J. Appl. Physiol.* 99, 330–337.
- Jing, E., Emanuelli, B., Hirschey, M.D., Boucher, J., Lee, K.Y., Lombard, D., Verdin, E.M., and Kahn, C.R. (2011). Sirtuin-3 (Sirt3) regulates skeletal muscle metabolism and insulin signaling via altered mitochondrial oxidation and reactive oxygen species production. *Proc. Natl. Acad. Sci. USA* 108, 14608–14613.
- Jing, E., O'Neill, B.T., Rardin, M.J., Kleinriders, A., Ilkayeva, O.R., Ussar, S., Bain, J.R., Lee, K.Y., Verdin, E.M., Newgard, C.B., et al. (2013). Sirt3 regulates metabolic flexibility of skeletal muscle through reversible enzymatic deacetylation. *Diabetes* 62, 3404–3417.
- Kelley, D.E., Goodpaster, B., Wing, R.R., and Simoneau, J.A. (1999). Skeletal muscle fatty acid metabolism in association with insulin resistance, obesity, and weight loss. *Am. J. Physiol.* 277, E1130–E1141.
- Koves, T.R., Ussher, J.R., Noland, R.C., Slentz, D., Mosedale, M., Ilkayeva, O., Bain, J., Stevens, R., Dyck, J.R., Newgard, C.B., et al. (2008). Mitochondrial overload and incomplete fatty acid oxidation contribute to skeletal muscle insulin resistance. *Cell Metab.* 7, 45–56.
- Koves, T.R., Sparks, L.M., Kovalik, J.P., Mosedale, M., Arumugam, R., DeBalsi, K.L., Everingham, K., Thorne, L., Phielix, E., Meex, R.C., et al. (2013). PPAR γ coactivator-1 α contributes to exercise-induced regulation of intramuscular lipid droplet programming in mice and humans. *J. Lipid Res.* 54, 522–534.
- Lee, S., Rivera-Vega, M., Alsayed, H.M., Boesch, C., and Libman, I. (2014). Metabolic inflexibility and insulin resistance in obese adolescents with non-alcoholic fatty liver disease. *Pediatr. Diabetes*. Published online April 23, 2014. <http://dx.doi.org/10.1111/pedi.12141>.
- Li, L.O., Grevengoed, T.J., Paul, D.S., Ilkayeva, O., Koves, T.R., Pascual, F., Newgard, C.B., Muoio, D.M., and Coleman, R.A. (2014). Compartmentalized acyl-CoA metabolism in skeletal muscle regulates systemic glucose homeostasis. *Diabetes*. Published online July 28, 2014. <http://dx.doi.org/10.2337/db13-1070>.
- Lindeboom, L., Nabuurs, C.I., Hoeks, J., Brouwers, B., Phielix, E., Kooi, M.E., Hesselink, M.K., Wildberger, J.E., Stevens, R.D., Koves, T., et al. (2014). Long-echo time MR spectroscopy for skeletal muscle acetylcarnitine detection. *J. Clin. Invest.* 124, 4915–4925.
- Liu, T.F., Vachharajani, V.T., Yoza, B.K., and McCall, C.E. (2012). NAD⁺-dependent sirtuin 1 and 6 proteins coordinate a switch from glucose to fatty acid oxidation during the acute inflammatory response. *J. Biol. Chem.* 287, 25758–25769.
- Mailloux, R.J., Jin, X., and Willmore, W.G. (2013). Redox regulation of mitochondrial function with emphasis on cysteine oxidation reactions. *Redox Biol.* 2, 123–139.
- McGarry, J.D. (2002). Banting lecture 2001: dysregulation of fatty acid metabolism in the etiology of type 2 diabetes. *Diabetes* 51, 7–18.
- Muoio, D.M., and Neuffer, P.D. (2012). Lipid-induced mitochondrial stress and insulin action in muscle. *Cell Metab.* 15, 595–605.
- Muoio, D.M., and Newgard, C.B. (2008). Mechanisms of disease: molecular and metabolic mechanisms of insulin resistance and beta-cell failure in type 2 diabetes. *Nat. Rev. Mol. Cell Biol.* 9, 193–205.
- Muoio, D.M., Noland, R.C., Kovalik, J.P., Seiler, S.E., Davies, M.N., DeBalsi, K.L., Ilkayeva, O.R., Stevens, R.D., Kheterpal, I., Zhang, J., et al. (2012). Muscle-specific deletion of carnitine acetyltransferase compromises glucose tolerance and metabolic flexibility. *Cell Metab.* 15, 764–777.
- Newgard, C.B. (2012). Interplay between lipids and branched-chain amino acids in development of insulin resistance. *Cell Metab.* 15, 606–614.
- Newman, J.C., He, W., and Verdin, E. (2012). Mitochondrial protein acylation and intermediary metabolism: regulation by sirtuins and implications for metabolic disease. *J. Biol. Chem.* 287, 42436–42443.
- Noland, R.C., Koves, T.R., Seiler, S.E., Lum, H., Lust, R.M., Ilkayeva, O., Stevens, R.D., Hegardt, F.G., and Muoio, D.M. (2009). Carnitine insufficiency caused by aging and overnutrition compromises mitochondrial performance and metabolic control. *J. Biol. Chem.* 284, 22840–22852.
- Papa, L., and Germain, D. (2014). Sirt3 regulates the mitochondrial unfolded protein response. *Mol. Cell Biol.* 34, 699–710.
- Peek, C.B., Affinati, A.H., Ramsey, K.M., Kuo, H.Y., Yu, W., Sena, L.A., Ilkayeva, O., Marcheva, B., Kobayashi, Y., Omura, C., et al. (2013). Circadian clock NAD⁺ cycle drives mitochondrial oxidative metabolism in mice. *Science* 342, 1243417.
- Prior, S.J., Ryan, A.S., Stevenson, T.G., and Goldberg, A.P. (2014). Metabolic inflexibility during submaximal aerobic exercise is associated with glucose intolerance in obese older adults. *Obesity (Silver Spring)* 22, 451–457.
- Ramsay, R.R., and Zammit, V.A. (2004). Carnitine acyltransferases and their influence on CoA pools in health and disease. *Mol. Aspects Med.* 25, 475–493.

- Randle, P.J. (1998). Regulatory interactions between lipids and carbohydrates: the glucose fatty acid cycle after 35 years. *Diabetes Metab. Rev.* *14*, 263–283.
- Shimomura, Y., Obayashi, M., Murakami, T., and Harris, R.A. (2001). Regulation of branched-chain amino acid catabolism: nutritional and hormonal regulation of activity and expression of the branched-chain alpha-keto acid dehydrogenase kinase. *Curr. Opin. Clin. Nutr. Metab. Care* *4*, 419–423.
- Sparks, L.M., Ukropcova, B., Smith, J., Pasarica, M., Hymel, D., Xie, H., Bray, G.A., Miles, J.M., and Smith, S.R. (2009). Relation of adipose tissue to metabolic flexibility. *Diabetes Res. Clin. Pract.* *83*, 32–43.
- Sugden, M.C., and Holness, M.J. (2006). Mechanisms underlying regulation of the expression and activities of the mammalian pyruvate dehydrogenase kinases. *Arch. Physiol. Biochem.* *112*, 139–149.
- Thyfault, J.P. (2008). Setting the stage: possible mechanisms by which acute contraction restores insulin sensitivity in muscle. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* *294*, R1103–R1110.
- Trepanowski, J.F., Canale, R.E., Marshall, K.E., Kabir, M.M., and Bloomer, R.J. (2011). Impact of caloric and dietary restriction regimens on markers of health and longevity in humans and animals: a summary of available findings. *Nutr. J.* *10*, 107.
- Turer, A.T., Malloy, C.R., Newgard, C.B., and Podgoreanu, M.V. (2010). Energetics and metabolism in the failing heart: important but poorly understood. *Curr. Opin. Clin. Nutr. Metab. Care* *13*, 458–465.
- Ukropcova, B., Sereda, O., de Jonge, L., Bogacka, I., Nguyen, T., Xie, H., Bray, G.A., and Smith, S.R. (2007). Family history of diabetes links impaired substrate switching and reduced mitochondrial content in skeletal muscle. *Diabetes* *56*, 720–727.
- Vadvalkar, S.S., Baily, C.N., Matsuzaki, S., West, M., Tesiram, Y.A., and Humphries, K.M. (2013). Metabolic inflexibility and protein lysine acetylation in heart mitochondria of a chronic model of type 1 diabetes. *Biochem. J.* *449*, 253–261.
- van Dijk, J.W., Venema, M., van Mechelen, W., Stehouwer, C.D., Hartgens, F., and van Loon, L.J. (2013). Effect of moderate-intensity exercise versus activities of daily living on 24-hour blood glucose homeostasis in male patients with type 2 diabetes. *Diabetes Care* *36*, 3448–3453.
- Wagner, G.R., and Payne, R.M. (2013). Widespread and enzyme-independent N ϵ -acetylation and N ϵ -succinylation of proteins in the chemical conditions of the mitochondrial matrix. *J. Biol. Chem.* *288*, 29036–29045.
- Wasserman, D.H., Kang, L., Ayala, J.E., Fueger, P.T., and Lee-Young, R.S. (2011). The physiological regulation of glucose flux into muscle in vivo. *J. Exp. Biol.* *214*, 254–262.

Metabolic Control of Autophagy

Lorenzo Galluzzi,^{1,2,3,4,5} Federico Pietrocola,^{1,2,3,4,5,6} Beth Levine,^{7,8} and Guido Kroemer^{1,2,3,4,9,10,*}

¹Equipe 11 labellisée par la Ligue Nationale contre le Cancer, Centre de Recherche des Cordeliers, 75006 Paris, France

²INSERM, U1138, 75006 Paris, France

³Université Paris Descartes, Sorbonne Paris Cité, 75005 Paris, France

⁴Université Pierre et Marie Curie, 75005 Paris, France

⁵Gustave Roussy Cancer Campus, 94805 Villejuif, France

⁶Université Paris Sud, 94805 Villejuif, France

⁷Center for Autophagy Research, Department of Internal Medicine, Department of Microbiology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

⁸Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

⁹Pôle de Biologie, Hôpital Européen Georges Pompidou, AP-HP, 75015 Paris, France

¹⁰Metabolomics and Cell Biology Platforms, Gustave Roussy Cancer Campus, 94805 Villejuif, France

*Correspondence: kroemer@orange.fr

<http://dx.doi.org/10.1016/j.cell.2014.11.006>

Macroautophagy (herein referred to as autophagy) is an evolutionarily conserved mechanism of adaptation to adverse microenvironmental conditions, including limited nutrient supplies. Several sensors interacting with the autophagic machinery have evolved to detect fluctuations in key metabolic parameters. The signal transduction cascades operating downstream of these sensors are highly interconnected to control a spatially and chronologically coordinated autophagic response that maintains the health and function of individual cells while preserving organismal homeostasis. Here, we discuss the physiological regulation of autophagy by metabolic circuitries, as well as alterations of such control in disease.

Introduction

Macroautophagy (hereafter referred to as autophagy) involves the sequestration of cytoplasmic components (which can be entire organelles, lipid vesicles, or protein aggregates) within a double-membraned vesicle, the so-called autophagosome. Autophagosomes fuse with lysosomes to generate autolysosomes, in which the autophagic cargo is degraded by acidic hydrolases. Autophagy relies on a machinery that operates in a tightly coordinated fashion and includes: (1) a multiprotein complex organized around unc-51-like autophagy activating kinase 1 (ULK1), RB1-inducible coiled-coil 1 (RB1CC1, best known as FIP200), autophagy-related 13 (ATG13), and ATG101, which triggers autophagy when the mechanistic target of rapamycin (mTOR) complex 1 (mTORC1) is inhibited; (2) a second multiprotein complex involving (among several interactors) phosphatidylinositol 3-kinase, catalytic subunit type 3 (PIK3C3, best known as vacuolar protein sorting 34, VPS34), Beclin 1 (BECN1), and autophagy/beclin-1 regulator 1 (AMBRA1), which favors the nucleation of autophagosome precursors (so-called isolation membranes or phagophores) when inhibitory signals from antiapoptotic members of the Bcl-2 protein family are blocked; (3) two transmembrane proteins, ATG9 and vacuole membrane protein 1 (VMP1), which recycle between the Golgi apparatus, endosomes, and autophagosomes, probably facilitating the recruitment of lipids to isolation membranes; (4) two ubiquitin-like (UBL) protein conjugation systems, which cooperate to catalyze the covalent attachment of ATG12 to ATG5 and ATG16-like 1 (ATG16L1) and that of phosphatidylethanolamine to microtubule-associated protein 1 light chain 3 (MAP1LC3, best known

as LC3); (5) several soluble NSF attachment protein receptor (SNARE)-like proteins, which promote the fusion between autophagosomes and lysosomes; and (6) various lysosomal enzymes that hydrolyze complex carbohydrates, proteins, lipids, and nucleic acids at low pH (for review, see Mizushima [2007]).

The primary, phylogenetically conserved role of autophagy is presumably to maintain cellular homeostasis in conditions of dwindling nutrient supplies and other metabolic perturbations (e.g., hypoxia). This is achieved through the rapid mobilization of endogenous reserves, aimed at retrieving fuel for ATP synthesis as well as building blocks for essential anabolic reactions (Singh and Cuervo, 2011), coupled to a global rewiring of intracellular metabolism (Figure 1). Autophagy-deficient eukaryotic cells are more sensitive to nutrient deprivation than their wild-type counterparts (Kroemer et al., 2010), and established tumors may be addicted to autophagy as a means to cope with adverse microenvironmental conditions (Guo et al., 2013a). Moreover, mice with genetic defects in essential components of the autophagic machinery die shortly after birth partly because they fail to mobilize sufficient reserves to survive the period of starvation between placental metabolism and breast feeding (Kuma et al., 2004).

Autophagy can be relatively nonselective, targeting to lysosomal degradation virtually any portion of the cytoplasm, or it may dispose of specific subcellular compartments in a highly selective manner (Mizushima and Komatsu, 2011). Generally, autophagic responses triggered by nutrient deprivation (which mainly serve bioenergetic/metabolic functions) are of the former type, although elongated mitochondria are selectively spared from

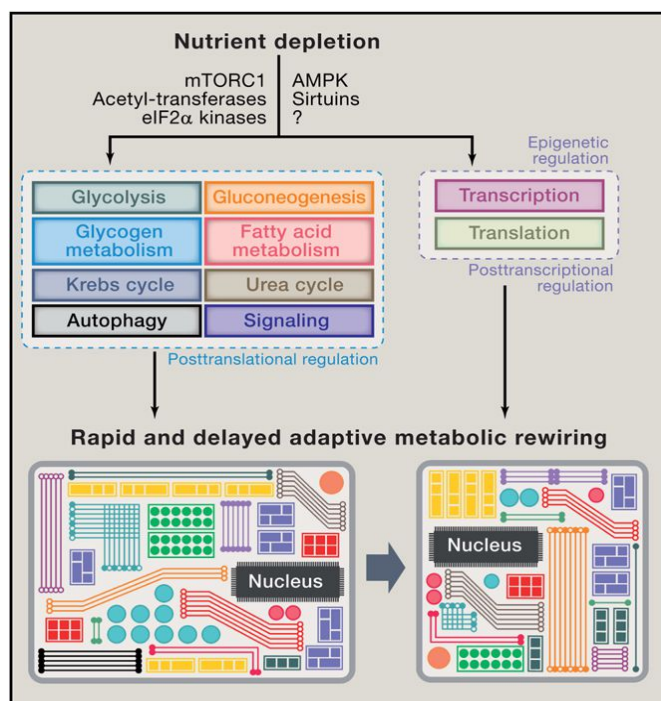


Figure 1. Cell-wide Metabolic Rewiring Associated with the Activation of Autophagy

In response to several perturbations of homeostasis, including declining levels of nutrients, cells mount an adaptive response organized around the autophagy-dependent mobilization of intracellular reserves. This response is biphasic, as it involves rapid posttranslational modifications as well as a transcriptional and translational reprogramming that has delayed consequences. Moreover, it is accompanied by a cell-wide rewiring of multiple metabolic circuitries, including both catabolic and anabolic pathways, which sustains cell survival and ensures basic cellular functions in conditions of stress. AMPK, 5' AMP-activated protein kinase; eIF2 α , eukaryotic translation initiation factor 2 α ; mTORC1, mechanistic target of rapamycin complex 1.

degradation in this context. Conversely, organellar damage or intracellular pathogens trigger highly selective forms of autophagy (Mizushima and Komatsu, 2011). Of note, autophagy can also actively participate in both programmed and stress-induced instances of cell death (Galluzzi et al., 2014), but this aspect will not be discussed further here.

Autophagy is crucial not only for adaptive responses to stress, but also for the maintenance of cellular homeostasis in physiological settings, at least in part because it mediates the removal of potentially dangerous constituents such as protein aggregates and dysfunctional mitochondria (Green et al., 2011). In line with this notion, the activation of autophagy at the whole-body level extends the lifespan of various model organisms, including mice (Rubinsztein et al., 2011). Moreover, defects in the autophagic machinery have been associated with numerous diseases, including aging-associated pathologies, neurodegeneration, cancer, cardiovascular disorders, and infectious/inflammatory conditions, as well as metabolic problems (Table S1 available online) (Choi et al., 2013). Thus, autophagy-incompetent mice develop both genetically and chemically driven neoplasms at a higher incidence than their autophagy-competent counterparts (Guo et al., 2013a). Various experimental models

of obesity and insulin resistance are also characterized by defects in hepatic autophagy that can be efficiently targeted to prevent steatosis/steatohepatitis and improve insulin sensitivity (Yang et al., 2010). This is not surprising, given the central position occupied by the liver in the regulation of organismal metabolism and the role of autophagy in the rewiring of intracellular metabolic circuitries.

Here, we will discuss the intimate crosstalk between metabolism and autophagy, placing special emphasis on the mechanisms through which nutrients and metabolic byproducts induce or suppress autophagy at the single-cell and whole-body level, and we will explore how the metabolic regulation of autophagy influences organismal fitness in health and disease.

Metabolic Triggers of Autophagy

In isolated cells, autophagy is generally induced by limitations in ATP availability or a lack of essential nutrients, including glucose and amino acids, yet it can also be stimulated by the accumulation of specific metabolites or metabolic byproducts, such as fatty acids and ammonia (Figure 2).

Reduced Energy Charge

The metabolic status of a cell can be represented by the “energy charge” of the adenylate system (a function of intracellular ATP, ADP, and AMP concentrations), which is calculated according to the formula $([ATP] + 1/2 [ADP])/([ATP] + [ADP] + [AMP])$ (Atkinson and Walton, 1967). When ATP is not actively synthesized through glycolysis or oxidative phosphorylation, the energy charge decreases in parallel with the accumulation of AMP, a condition that stimulates autophagy through protein kinase, AMP-activated (PRKA, best known as 5' AMP-activated protein kinase, AMPK) (Hardie et al., 2012). Because AMPK utilizes ATP as a donor of phosphate groups and because several steps in the autophagic cascade consume energy, a minimum amount of ATP is required for the induction of autophagy. Thus, a rapid reduction of the energy charge below a critical limit is likely to trigger cell death rather than an adaptive autophagic response (Galluzzi et al., 2014). In cells that mostly rely on glycolysis, withdrawing glucose promotes autophagy as a result of AMP accumulation and the consequent activation of AMPK (Hardie et al., 2012). However, the inhibition of hexokinase 2 (HK2, the enzyme that catalyzes the first, rate-limiting step of glycolysis) with 2-deoxyglucose does not have the same effect because HK2 directly promotes autophagy by physically interacting and hence inhibiting mTORC1 (Roberts et al., 2014). Similarly, rotenone, a widely employed inhibitor of the respiratory chain, inhibits mitochondrial ATP synthesis but paradoxically inhibits autophagic flux (Mader et al., 2012). Thus, using toxins may not be an appropriate approach to probe complex circuitries such as those linking metabolism and autophagy. Of note, starvation, as well as hypoxia, are generally associated with increased amounts of reactive oxygen species (ROS). ROS promote autophagy by several mechanisms, including: (1) the hypoxia-inducible factor 1 (HIF-1)-dependent transactivation of BCL2/adenovirus E1B 19 kDa interacting protein 3 (BNIP3) and BNIP3-like (BNIP3L), encoding two Bcl-2 family members that potently stimulate the autophagic removal of dysfunctional mitochondria (mitophagy); (2) the ataxia-telangiectasia mutated (ATM)-dependent activation of tuberous sclerosis 2 (TSC2), a major suppressor of

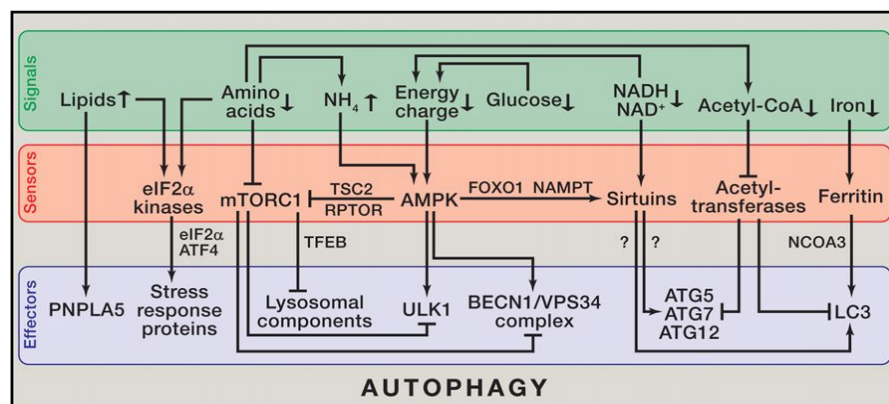


Figure 2. Metabolic Regulation of Autophagy at the Single-Cell Level

Several changes in the availability of nutrients in extracellular fluids trigger autophagy (directly or indirectly), including drops in the levels of glucose, amino acids, acetyl-CoA, and iron and decreases in the relative abundance reduced versus oxidized NAD, as well as the accumulation of specific lipids and ammonia (NH₄, a product of amino acid catabolism). Many of these metabolic cues stimulate autophagy because they inhibit mechanistic target of rapamycin complex 1 (mTORC1) or various acetyltransferases or because they activate 5' AMP-activated protein kinase (AMPK), deacetylases of the sirtuin family, or eukaryotic translation initiation factor 2α (eIF2α) kinases. In addition, the accumulation of lipids may directly favor the formation of autophagosomes in a patatin-like phospholipase-domain-containing 5 (PNPLA5)-dependent fashion, whereas the

depletion of iron has autophagy-stimulating effects upon the recognition of ferritin heavy and light chains by the autophagic adaptor nuclear receptor coactivator 4 (NCOA4). Both mTORC1 and AMPK regulate autophagy by controlling the activity of essential components of the autophagic machinery, such as unc-51 like autophagy activating kinase 1 (ULK1) or Beclin 1 (BECN1). Moreover, they are both involved in mutually regulatory interactions as well as in functional interactions with other nutrient sensors, such as sirtuins. ATF4, activating transcription factor 4; ATG, autophagy-related; FOXO1, forkhead box O1; NAMPT, nicotinamide phosphoribosyltransferase; RPTOR, regulatory-associated protein of MTOR, complex 1; TFEB, transcription factor EB; TSC2, tuberous sclerosis 2; VPS34, vacuolar protein sorting 34.

mTORC1 signaling; and (3) the oxidation-dependent activation of the essential autophagic protein ATG4 (Scherz-Shouval and Elazar, 2011). In summary, glucose deprivation and the consequent alterations in energy charge and ROS levels are potent activators of autophagy.

Reduced NADH/NAD⁺ Ratio

In either its oxidized (NAD⁺) or reduced (NADH) form, NAD is an essential substrate for multiple metabolic circuitries, including (but not limited to) glycolysis, the Krebs cycle, and oxidative phosphorylation. The exposure of cells to nutrient-free conditions causes the accumulation of NAD⁺ at the expense of NADH, promoting autophagy upon activation of histone deacetylases of the sirtuin family (Houtkooper et al., 2012). Conversely, the intracellular levels of both NAD⁺ and NADH fall upon the activation of NAD⁺-dependent enzymes such as poly(ADP-ribose) polymerase 1 (PARP1) (Gibson and Kraus, 2012). Inhibition of these enzymes (which preserves the endogenous levels of NAD) as well as the artificial supply of NAD precursors (e.g., nicotinamide, nicotinamide riboside) potentially triggers autophagy upon the activation of sirtuins, whose enzymatic activity critically relies on NAD⁺ (Houtkooper et al., 2012). Thus, not only the relative abundance of NADH and NAD⁺, but also the total availability of NAD has profound autophagy-modulatory effects.

Depletion of Cytosolic Acetyl-CoA

The exposure of mammalian cells to nutrient-free conditions for several hours or the overnight starvation of mice causes a significant decrease in the cytosolic levels of acetyl-CoA, which correlates with the induction of autophagy (Mariño et al., 2014). A similar effect is observed with several pharmacological or genetic interventions that inhibit (directly or by limiting substrate availability) the synthesis of acetyl-CoA within mitochondria or in the cytosol (Eisenberg et al., 2014; Mariño et al., 2014). The depletion of cytosolic acetyl-CoA stores potentially stimulates autophagy, presumably because acetyl-CoA is the sole donor of acetyl groups for acetyl transferases, some of which regulate the activity of various components of the autophagic machinery

at the posttranslational level (Mariño et al., 2014) or their synthesis (by acetylating histones) (Lee et al., 2014). Conversely, when intracellular acetyl-CoA levels are replenished artificially, starvation-induced autophagy is inhibited both in vitro in cultured cells and in vivo in mice (Mariño et al., 2014). Of note, the constitutive activation of v-akt murine thymoma viral oncogene homolog 1 (AKT1, also known as PKB), which can be triggered by oncogenic alterations such as activating mutations in Kirsten rat sarcoma viral oncogene homolog (KRAS), promotes acetyl-CoA synthesis upon the phosphorylation-dependent activation of ATP citrate lyase (ACLY) (Lee et al., 2014). The consequent inhibition of autophagy may contribute, at least in part, to the oncogenic effects of AKT1 hyperactivation.

Depletion of Amino Acids

Limitations in the availability of nonessential amino acids can trigger autophagy through at least four distinct, nonmutually exclusive mechanisms. First, a drop in the intracellular abundance of amino acid results in the accumulation of uncharged tRNA species. This activates eukaryotic translation initiation factor 2α kinase 4 (EIF2AK4, best known as GCN2), which blocks protein synthesis and triggers autophagy via activating transcription factor 4 (ATF4) (Ye et al., 2010). Second, the absence of amino acids in the lysosomal lumen turns off an “inside-outside” mechanism that promotes the recruitment of mTORC1 at the lysosomal surface and its activation (Zoncu et al., 2011). Third, the lack of various amino acids—in particular, leucine, glutamate, and glutamine—negatively affects intracellular acetyl-CoA stores (Mariño et al., 2014), reflecting the ability of these amino acids to efficiently feed into the Krebs cycle to generate acetyl-CoA. Fourth, the depletion of the key metabolic intermediate α-ketoglutarate caused by dwindling amino acid levels promotes autophagy along with the inhibition of proline hydroxylases (but not the stabilization of HIF-1) (Durán et al., 2013), reflecting the role of α-ketoglutarate as an obligate donor of hydroxyl groups for this class of enzymes. Proteasome inhibitors can also cause a drop in the intracellular availability of amino

acids and hence trigger autophagy. Although the relative weight of each of these pathways has not yet been determined, it appears plausible that all of these mechanisms contribute to the orchestration of optimal autophagic responses upon amino acid shortage.

Depletion of Iron

Iron is an obligate cofactor for several enzymes that catalyze redox reactions, including components of the mitochondrial respiratory chain. A fraction of cytoplasmic iron is stored within large ferritin oligomers, which can be rapidly degraded by the autophagic machinery to serve cellular needs. Drops in the intracellular availability of free iron (which can be mimicked by the administration of pharmacological chelators) activate an autophagic response that has been termed “ferritinophagy.” This appears to require nuclear receptor coactivator 4 (NCOA4), which operates as an autophagic receptor for the recognition and engulfment of ferritin light and heavy chains by LC3-containing autophagosomes (Mancias et al., 2014). This example illustrates how autophagy can help cells to palliate a selective micronutrient deficiency. It will be important to explore whether similar mechanisms exist for specifically mobilizing other nutrients.

Increased Ammonia Levels

Ammonia is one of the main byproducts of the catabolism of amino acids and a potent inducer of autophagy (Eng et al., 2010). However, in contrast to the autophagic response to decreased amino acid availability, ammonia-induced autophagy does not rely on ULK1/ULK2 activation (Cheong et al., 2011) or mTORC1 inhibition (Harder et al., 2014). Rather, it seems that ammonia triggers autophagy by activating AMPK and favoring the endoplasmic reticulum (ER) stress response (Harder et al., 2014). Of note, neoplastic tissues produce high levels of ammonia as a result of an intense flux through glutaminolysis (Galluzzi et al., 2013). At least in part, this may contribute to the upregulation of the autophagic flux observed in some established neoplasms (Guo et al., 2013a).

Lipids

Both saturated and unsaturated fatty acids, such as palmitate and oleate, respectively, can stimulate autophagy, albeit through distinct mechanisms. Palmitate-induced, but not oleate-induced, autophagy requires EIF2AK2 (best known as PKR) and mitogen-activated protein kinase 8 (MAPK8, best known as JNK1) (Shen et al., 2012). Stearoyl-CoA desaturase, which converts saturated lipids into their monounsaturated counterparts, is required for starvation-induced autophagy, and the external supplementation of oleate can overcome the autophagic defect induced by stearoyl-CoA desaturase inhibitors (Ogasawara et al., 2014). Possibly, this results from the need for lipids in the generation of autophagosomes, a process that may rely on the neutral lipase patatin-like phospholipase-domain-containing 5 (PNPLA5) (Dupont et al., 2014). Enterocytes transiently store dietary lipids in triglyceride-containing droplets that localize at the ER. Such droplets trigger an immediate autophagic response that results in their capture by nascent autophagosomes and their delivery to lysosomes for degradation (Khalidoun et al., 2014). Hence, lipids can induce autophagy despite being nutrients, and this may constitute an important mechanism to avoid lipotoxicity at the cell-autonomous level.

Metabolic Sensors that Initiate Autophagy **AMPK**

AMPK, one of the key energy sensors of the cell, is a heterotrimer composed of a catalytic α subunit, a scaffolding β subunit, and a regulatory γ subunit, all of which are expressed in multiple variants by mammals ($\alpha 1$, $\alpha 2$, $\beta 1$, $\beta 2$, $\gamma 1$, $\gamma 2$, and $\gamma 3$). The binding of two molecules of AMP (or ADP, with lower affinity) to the γ subunit inhibits the inactivating dephosphorylation of the α subunit at T172 (Hardie et al., 2012). Hence, decreases in cellular energy charge boost the kinase activity of AMPK. The phosphorylation of AMPK α subunit at T172 can be catalyzed by calcium/calmodulin-dependent protein kinase kinase 2, β (CAMKK2) and serine/threonine kinase 11 (STK11, best known as liver kinase B1, LKB1), or stimulated (probably via an indirect mechanism) by mitogen-activated protein kinase kinase 7 (MAP3K7, best known as TAK1) (Hardie et al., 2012).

TAK1 is required for the starvation-induced phosphorylation of AMPK and consequent autophagic response in cancer cells in vitro (Criollo et al., 2011), as well as in hepatocytes in vivo (Inokuchi-Shimizu et al., 2014). TAK1 activation is linked to two additional phenomena that may stimulate autophagy, namely: (1) the displacement of TAK1-binding protein 2 (TAB2) and TAB3, two TAK1 coactivators, from their autophagy-inhibitory interaction with BECN1 (Criollo et al., 2011), and (2) the activation of the I κ B kinase (IKK) complex, which stimulates autophagy by phosphorylating the regulatory subunit of phosphoinositide-3-kinase (PI3K), thereby reducing its localization to cell membranes and enzymatic activity (Comb et al., 2012). Indeed, IKK is required for an optimal autophagic response to starvation in vitro (Comb et al., 2012; Criollo et al., 2011), and the ablation of the gene coding for the IKK-subunit-conserved helix-loop-helix ubiquitous kinase (CHUK, best known as IKK α), limits autophagy in pancreatic acinar cells in vivo (Li et al., 2013). Under some circumstances, AMPK can activate TAK1 (Lanna et al., 2014), suggesting that these kinases may engage in a mutually stimulatory amplification cascade. AMPK can also be activated allosterically by pharmacological agents, perhaps reflecting the existence of a hitherto unidentified (and perhaps crucial) endogenous metabolite that regulates its enzymatic activity. Among other compounds, this applies to salicylate, a phenolic phyto-hormone with analgesic, antipyretic, anti-inflammatory, and perhaps anticancer activity (Hawley et al., 2012). Still, it remains unclear whether the activation of AMPK is relevant to the broad pharmacological effects of aspirin, the pro-drug of salicylate.

AMPK is a master regulator of metabolism, and it stimulates autophagy by multiple mechanisms. Beyond inhibiting mTORC1 (see below), AMPK phosphorylates and activates ULK1 (Kim et al., 2011), as well as various components of the BECN1/VPS34 complex. In particular, upon glucose deprivation, AMPK phosphorylates BECN1 on S93 and S96, which augments the class III PI3K activity of VPS34, as well as VPS34 itself (on T163 and S165), which inhibits its nonautophagic functions in endosome-to-Golgi retrograde trafficking (Kim et al., 2013).

mTORC1

mTORC1 is composed of: (1) MTOR; (2) two mTORC1-specific proteins—namely, regulatory-associated protein of MTOR, complex 1 (RPTOR) and AKT1 substrate 1 (AKT1S1, best known as PRAS40); and (3) several proteins that are shared with another

MTOR-containing complex (mTORC2), i.e., DEP-domain-containing MTOR-interacting protein (DEPTOR) and MTOR-associated protein, LST8 homolog (MLST8). In response to growth factors, mTORC1 phosphorylates eukaryotic translation initiation factor 4E binding protein 1 (EIF4EBP1, best known as 4-EBP1) and ribosomal protein S6 kinase (RPS6K, best known as p70^{S6K}), ultimately promoting protein synthesis (Shimobayashi and Hall, 2014). Activated mTORC1 suppresses autophagy by phosphorylating and inhibiting ULK1 (Kim et al., 2011), AMBRA1 and ATG14 (two autophagy-stimulatory interactors of BECN1) (Nazio et al., 2013; Yuan et al., 2013), and transcription factor EB (TFEB, see below) (Settembre et al., 2013). Of note, AMPK can inhibit mTORC1 (and hence promote autophagy) indirectly by phosphorylating and activating tuberous sclerosis 2 (a negative regulator of mTORC1), as well as directly by phosphorylating RPTOR (Hardie et al., 2012).

The availability of amino acids positively regulates mTORC1 (and hence suppresses autophagy) via multiple pathways. For instance, amino acids activate mTORC1 on the lysosomal membrane from within the lumen of the organelle, a process that involves vacuolar ATPases, a Ras-related GTP binding (RRAG)-containing complex that has been termed “Ragulator,” and the mTORC1 activator Ras homolog enriched in brain (RHEB) (Sancak et al., 2010). Artificial increases in the levels of α -ketoglutarate, which can be achieved by the provision of glutamine (via glutaminolysis) or several cell-permeant precursors (i.e., dimethyl- α -ketoglutarate, trifluoromethylbenzyl- α -ketoglutarate, 5-ethyltrifluoromethylbenzyl- α -ketoglutarate), potently activate mTORC1 in cells depleted of amino acids (Mariño et al., 2014), at least in part through such a lysosomal RHEB-dependent pathway (Durán et al., 2012). Conversely, another α -ketoglutarate precursor (i.e., 1-octyl- α -ketoglutarate) appears to inhibit mTORC1, activate autophagy, and mediate lifespan-extending effects (Chin et al., 2014). The authors of this report ascribe their findings to the ability of α -ketoglutarate to inhibit mitochondrial ATP synthesis at the level of the F₁F₀-ATPase. Of note, leucine may also activate mTORC1 through the lysosomal RRAG-RHEB system. Indeed, leucine has been shown to stimulate glutaminolysis by allosterically activating glutamate dehydrogenase and to activate RRAG in the form of leucyl-tRNA synthetase (Han et al., 2012). Recently, the lysosomal RRAG-RHEB system has been suggested to contribute to the activation of mTORC1 by glucose (Efeyan et al., 2013), suggesting that AMPK may not constitute the sole sensor of glucose deprivation.

It is important to note that mTORC1 not only represses autophagy and lysosomal biogenesis, but also operates as a general regulator of anabolic reactions (Shimobayashi and Hall, 2014). Thus, similar to AMPK, mTORC1 controls several metabolic circuitries outside of the autophagic cascade, implying that chemical mTORC1 inhibitors such as rapamycin (which is approved for use in humans as an immunosuppressant to prevent the rejection of solid transplants) and other compounds commonly referred to as “rapalogs” have broad metabolic consequences that are not limited to the induction of autophagy. Moreover, mTORC1 inhibitors lose their capacity to trigger autophagy when the downstream signaling pathways are affected by oncogenic autophagy-suppressing alterations, such as the phosphorylation of BECN1 on S234 and S295 (which is catalyzed

by AKT1) (Wang et al., 2012) or on Y229, Y233, and Y352, which is catalyzed by the epidermal growth factor receptor (EGFR) (Wei et al., 2013). This should be taken into consideration when mTORC1 inhibitors are employed to stimulate autophagy in cancer cells exhibiting PI3K hyperactivation or bearing activating mutations in *EGFR*.

elf2 α Kinases

The phosphorylation of eukaryotic translation initiation factor 2 α (EIF2A, best known as elf2 α) on S51 is a cardinal feature of the so-called “integrated stress response,” which allows cells to interrupt protein synthesis in response to the accumulation of unfolded proteins in the ER and attempt to restore homeostasis along with the activation of autophagy (Kroemer et al., 2010). The mammalian genome codes for at least four kinases that phosphorylate elf2 α —namely, EIF2AK1 (best known as HRI), EIF2AK2 (best known as PKR), EIF2AK3 (best known as PERK), and EIF2AK4 (best known as GCN2) (Silvera et al., 2010). These kinases are activated by a variety of stimuli. Limited heme availability or heavy metals like cadmium activate HRI. Double-stranded RNA or high doses of palmitate stimulate PKR. The accumulation of unfolded proteins in the ER activates PERK. Finally, the accumulation of uncharged tRNAs boosts the activity of GCN2 (Silvera et al., 2010). Of note, the response of cells expressing a nonphosphorylatable mutant of elf2 α (EIF2A^{S51A}) to several autophagy-inducing conditions is largely defective (Tallóczy et al., 2002). This may indicate that the efficient induction of autophagy by various stimuli requires the transactivation of multiple genes that are controlled by ATF4, which operates downstream of elf2 α (B’chir et al., 2013). Thus, it appears that the integrated stress response is closely tied to the regulation of autophagy.

Sirtuins

Sirtuins constitute a family of NAD⁺-dependent class III histone deacetylases that catalyze the deacetylation of protein substrates coupled to the generation of nicotinamide and 2'-O-acetyl-ADP-ribose (Houtkooper et al., 2012). The best-studied of these enzymes, sirtuin 1 (SIRT1), is mainly located in the nucleus, where it deacetylates various histones (e.g., H1, H3, and H4) and other proteins, including transcription factors such as p53, NF- κ B, forkhead box O1 (FOXO1), FOXO3, and peroxisome proliferator-activated receptor γ , coactivator 1 α (PPARGC1A) (Houtkooper et al., 2012). Activation of SIRT1 with the natural polyphenol resveratrol promotes autophagy and extends the lifespan of several organisms (Lagouge et al., 2006). Moreover, overexpression of a SIRT1 mutant that exclusively localizes to the cytoplasm induces robust autophagic responses (Morselli et al., 2011), indicating that SIRT1 can trigger autophagy independently of its transcriptional functions. The cytoplasmic effectors of SIRT1-driven autophagy remain to be precisely identified, although one single report points to a direct involvement of ATG5, ATG7, ATG12 and LC3 (Lee et al., 2008).

Besides responding to increasing NAD⁺ concentrations (see above), the enzymatic activity of SIRT1 changes as a function of its own expression levels. Although high-fat diet and obesity downregulate SIRT1 in several organs, both in mice and in humans, caloric restriction promotes SIRT1 expression in multiple mouse tissues (Chalkiadaki and Guarente, 2012). Because the ability of caloric restriction to increase lifespan is preserved in

Sir1^{+/-} (but lost in *Sir1*^{-/-}) mice (Mercken et al., 2014), such an upregulation may not be involved in the lifespan-extending effects of interventions that activate SIRT1. However, results obtained with *Sir1*^{-/-} mice must be taken with caution because these animals are not born at Mendelian ratios and often exhibit developmental defects (Cheng et al., 2003). Of note, AMPK may activate SIRT1 by promoting the FOXO1-dependent transactivation of the gene encoding nicotinamide phosphoribosyltransferase (NAMPT), an enzyme involved in the production of NAD⁺ (Cantó et al., 2009). Thus, at least under some circumstances, SIRT1 contributes to the proautophagic activity of AMPK.

Acetyltransferases

Reductions in the intracellular pool of acetyl-CoA entail a net decrease in global protein acetylation, both in the cytoplasm (Mariño et al., 2014) and in the nucleus (Eisenberg et al., 2014). Thus, the activity of several acetyltransferases may vary as a function of the availability of acetyl-CoA, the sole donor of acetyl groups for the reactions that they catalyze. One acetyltransferase that plays a critical role in the regulation of autophagy is E1A-binding protein p300 (EP300) (Mariño et al., 2014). In cell-free systems, the activity of EP300 responds to shifts in the intracellular abundance of acetyl-CoA observed during the physiological transition from a fed to an unfed state. Moreover, the genetic or pharmacological inhibition of EP300 promotes autophagy even in conditions in which acetyl-CoA is artificially maintained at high levels (Mariño et al., 2014). EP300 reportedly acetylates and inhibits several proteins of the core autophagic machinery such as ATG5, ATG7, ATG12, and LC3 (Lee and Finkel, 2009). However, it is likely that other acetyltransferases, including members of the inhibitor of growth (ING) family, participate in the regulation of autophagy via nuclear and cytoplasmic pathways (Mariño et al., 2014). In yeast, which lacks a bona fide EP300 ortholog, several histone acetyltransferases have been implicated in the transcriptional control of autophagy (Eisenberg et al., 2009). Moreover, at least in yeast, a specific combination of deacetylation- and acetylation-mimicking mutations in histone-coding genes can cause the constitutive overexpression of core components of the autophagic machinery, resulting in increased autophagic flux and lifespan extension (Eisenberg et al., 2014).

Distinct acetyltransferases are organized in a hierarchical manner so that the inhibition of one can be coupled to the activation of another. For example, EP300 acts as a negative regulator of α tubulin acetyltransferase 1 (ATAT1, also known as MEC17), which itself is a substrate of, and can be activated by, AMPK. The activity of EP300 is also negatively regulated by AMPK-dependent phosphorylation on S89 (Yang et al., 2001). Thus, in conditions of EP300 inhibition and AMPK activation (which are intimately linked), MEC17 promotes the hyperacetylation of α -tubulin, which has autophagy-stimulatory effects (Mackeh et al., 2014). In summary, both deacetylation and hyperacetylation events contribute to autophagy, and it would be an oversimplification to state that all acetyltransferases contribute to the repression of autophagy.

Transcription Factors

The so-called “coordinated lysosomal expression and regulation” (CLEAR) gene network, an ensemble of genes expressed in a synchronized manner in response to perturbations of lyso-

somal activity (most of which are relevant for autophagy), is activated by transcription factors, including TFEB and its homolog transcription factor E3 (TFE3) (Settembre et al., 2013). TFEB is recruited to lysosomal membranes by the Ragulator, allowing for its phosphorylation at S142 and S211. Phosphorylated TFEB is sequestered by chaperones of the 14-3-3 family, which actively prevent its translocation to the nucleus. Accordingly, the substitution of TFEB S142 and S211 with alanine residues results in its constitutive translocation to the nucleus. TFEB can also be phosphorylated on S142 by mitogen-activated protein kinase 1 (MAPK1, best known as ERK2), which indeed exerts autophagy-inhibitory functions (Settembre et al., 2013). The identity of the phosphatase that dephosphorylates TFEB at these residues remains to be elucidated.

Importantly, TFEB regulates its own transcription, implying the existence of a self-amplificatory signaling loop that perpetuates the autophagic response. How such a loop is turned off remains elusive. Another transcription factor, zinc finger with KRAB and SCAN domains 3 (ZKSCAN3), functionally antagonizes TFEB. Thus, in response to starvation or mTORC1 inhibition, ZKSCAN3 translocates from the nucleus to the cytoplasm, and the knockdown of ZKSCAN3 suffices to facilitate the induction of autophagy. Indeed, ZKSCAN3 represses the transcription of more than 60 TFEB target genes involved in autophagy and lysosomal functions (Settembre et al., 2013). Whether these transcription factors truly detect metabolic perturbations or whether they simply execute autophagic responses triggered by upstream sensors such as mTORC1 remains to be determined. Irrespective of this unknown, TFEB-induced autophagy has a central role in disease protection, as the viral delivery of a TFEB-coding construct to the liver prevents the hepatic accumulation of lipid vesicles in both diet-induced and genetic models of obesity (Settembre et al., 2013). Interestingly, the nematode ortholog of TFEB (HLH-30) is required for the induction of autophagy by longevity-extending manipulations in *Caenorhabditis elegans* (Lapierre et al., 2013).

Cell-Surface Nutrient Receptors

Several G protein-coupled receptors (GPCRs) expressed on the cell surface sample the extracellular microenvironment for nutrient availability and signal to the autophagic machinery. These include (but are not limited to): (1) G protein-coupled receptor, class C, group 6, member A (GPRC6A), γ -aminobutyric acid B receptor 1 (GABBR1), calcium-sensing receptor (CASR), heterodimeric taste receptors, and various metabotropic glutamate receptors, all of which sense one or more amino acids; (2) free fatty acid receptor 1 (FFAR1) and FFAR4, which detect long-chain fatty acids; and (3) FFAR2 and FFAR3, which are activated by short-chain fatty acids (Wauson et al., 2014). The signal transduction cascades linking each of these receptors to the autophagic machinery have not yet been precisely defined, but they all presumably operate by promoting increases in the intracellular levels of inositol-1,4,5-triphosphate and diacylglycerol, or those of cyclic AMP (Wauson et al., 2014).

In several cell types, including pancreatic β cells, cardiac myoblasts, and cervical carcinoma HeLa cells, the knockdown of either subunit of heterodimeric taste receptors—namely, taste receptor, type 1, member 1 (TAS1R1) and TAS1R3—promotes autophagy even in the presence of excess extracellular amino

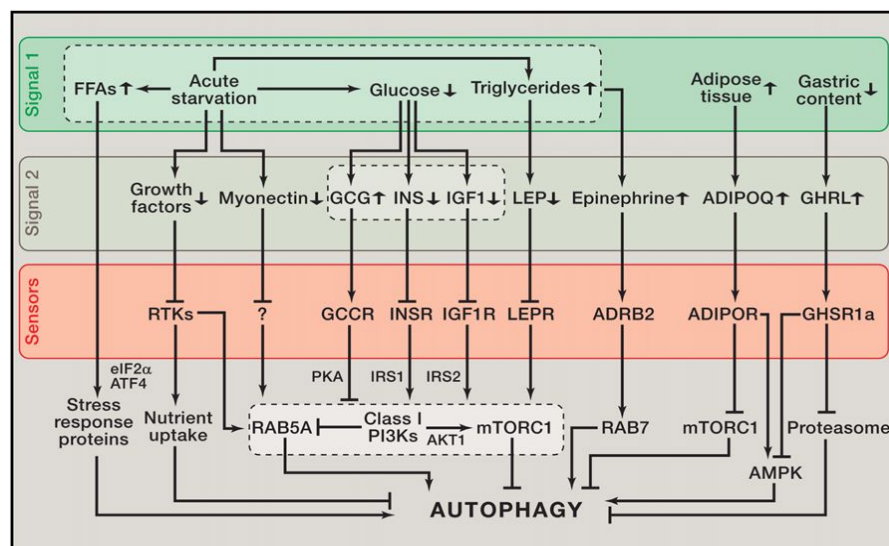


Figure 3. Metabolic Regulation of Autophagy at the Organismal Level

Acute starvation induces a stereotyped pattern of metabolic alterations, including a (limited) decrease in circulating glucose levels coupled to an increase in blood-borne triglycerides and free fatty acids (FFAs). This is generally accompanied by the secretion of glucagon (GCG), as well as by a reduction in the circulating levels of growth factors, insulin (INS), insulin-like growth factor 1 (IGF1), leptin (LEP), and myonectin. FFAs trigger autophagy as they freely enter cells and promote the inactivation of eukaryotic translation initiation factor 2 α (eIF2 α) coupled to the activation of activating transcription factor 4 (ATF4). Conversely, the alterations in the circulating levels of INS, IGF1, GCG, and LEP induced by starvation are sensed by specific receptors expressed at the cell surface, all of which impinge on the *v-akt* murine thymoma viral oncogene homolog 1 (AKT1)/mechanistic target of rapamycin complex 1 (mTORC1) signaling axis. In addition, drops in the availability of growth factors promote autophagy by limiting the expression of nutrient transporters. Epinephrine is also secreted in response to starvation, promoting autophagy in

the periphery upon binding to adrenoceptor β_2 (ADRB2). Other hormones with prominent autophagy-regulatory functions are adiponectin (ADIPOQ) and ghrelin (GHRL). The former, which is secreted by the adipose tissue, exerts pure autophagy stimulatory functions by inhibiting mTORC1 and promoting the activation of 5' AMP-activated protein kinase (AMPK). The latter, which is secreted upon relaxation of the gastric wall, has been shown to inhibit AMPK in some circumstances, hence inhibiting autophagy, and to suppress proteasomal protein degradation in others, hence increasing autophagic flux. ADIPOR, ADIPOQ receptor; GCGR, GCG receptor; GHRS1a, growth hormone secretagogue receptor 1A; IGF1R, IGF1 receptor; INSR, INS receptor; IRS, INSR substrate; LEP, LEP receptor; PI3K, phosphoinositide-3-kinase; PKA, protein kinase A; RTK, receptor tyrosine kinase.

acids. Similarly, *Tas1r3*^{-/-} mice exhibit increased autophagic responses to starvation in the heart, liver, and skeletal muscle, as compared to their wild-type counterparts (Wauson et al., 2012). Indirect evidence also suggests that the omega-3 fatty acid docosahexaenoic acid induces autophagy upon binding to FFAR4 on the cell surface (Williams-Bey et al., 2014). These results point to the possibility that GPCRs not only sense nutrients in the olfactory and gustative sensory organs, but also act in peripheral tissues to regulate autophagy in response to extracellular metabolic cues.

In summary, several systems are in place to detect fluctuations in the intracellular and/or extracellular availability of nutrients and hence initiate an autophagic response. However, it remains to be explored which among these systems preferentially respond to a global nutrient limitation (which may be caused by a reduction in blood supply) rather than to changes in the abundance of a specific molecule (which may be the result of precise metabolic perturbations). Moreover, it is not yet known whether nutrient sensors have similar activation thresholds in all cell types. It is reasonable to expect that distinct AMPK isoforms, the composition of mTORC1, the subcellular localization of acetyltransferases and sirtuins, as well as the expression pattern of GPCRs, ultimately impact on the fine regulation of autophagy.

Induction of Autophagy by Metabolic Restriction In Vivo

Although culturing cells in the absence of glucose, amino acids, or all nutrients constitutes a valid model for the induction of autophagy in vitro, such drastic alterations in the abundance of extracellular supplies do not occur in vivo, at least in mammals. Maintaining mice for 24–48 hr in the absence of food (but with free access to water) induces autophagy in close-to-all nucle-

ated cells of the body while causing a reduction of 10%–20% in body weight (Mizushima, 2009). Yet, this does not cause a major depletion in the circulating levels of amino acids or massive, life-threatening hypoglycemia because of the autophagy-dependent mobilization of cellular stores and the systemic response to starvation involving hepatic and muscular reserves (He et al., 2012; Kuma et al., 2004). Indeed, in multicellular eukaryotes, the composition of the extracellular milieu is preserved by multiple homeostatic circuits. Moreover, the cellular availability of nutrients is not mainly dictated by their abundance but, rather, by the regulation of their uptake via specific transporters expressed on the plasma membrane (Wieman et al., 2007). Thus, the expression levels and activity of the cellular systems that ensure the uptake of various nutrients, including glucose and amino acids, are regulated by several growth factors as well as by neuroendocrine circuits (Kim and Lee, 2014) (Figure 3).

In conditions of acute starvation, the circulating levels of insulin (INS) and insulin-like growth factor 1 (IGF1) decrease while those of the insulin-like growth factor binding protein 1 (IGFBP1, an IGF1 antagonist) and glucagon (GCG) increase (Cheng et al., 2014). The consequent reduction of INS and IGF1 signaling may contribute to the inhibition of nutrient uptake and mTORC1 inactivation, favoring a compensatory autophagic response that preserves bioenergetic homeostasis (Troncoso et al., 2012). Similarly, the absence of growth factors not only limits glucose uptake upon the downregulation of plasma membrane transporters (Wieman et al., 2007), but also inhibits downstream signaling via the AKT1/mTORC1 pathway and promotes the interaction between the catalytic subunit of class I PI3Ks and the small GTPase RAB5A, hence favoring the activating interaction of the latter with VPS34 (Dou et al., 2013). A similar

suppression of AKT1/mTORC1 signaling occurs upon the activation of protein kinase A by GCG (Kondomerkos et al., 2005). Thus, drops in the extracellular availability of INS, IGF1, and growth factors coupled to increased GCG signaling provoke a robust autophagic response. Accordingly, the postnatal increase in circulating INS levels resulting from breast feeding suppresses maladaptive autophagy in cardiomyocytes, as it has been shown in mice harboring a cardiomyocyte-specific deletion of *Irs1* and *Irs2* (coding for two key transducers of INS and IGF1 signals) (Riehle et al., 2013). Moreover, the exogenous provision of IGF1 reverts some of the metabolic effects of starvation in mice (Cheng et al., 2014).

Epinephrine is secreted by adrenal glands when hypothalamic neurons detect a drop in circulating glucose levels, and the consequent activation of β -adrenergic GPCRs in peripheral tissues promotes the mobilization of triglyceride stores through a mechanism that involves autophagy (Lizaso et al., 2013). Moreover, epinephrine deficiency (owing to the ablation of the gene coding for phenylethanolamine N-methyltransferase) causes severe hepatic steatosis coupled to deficient autophagy and impaired triglyceride usage yet does not affect glucose homeostasis (Sharara-Chami et al., 2012). Intriguingly, starvation is coupled to a major increase in circulating triglycerides, and the free fatty acids resulting from their catabolism (such as oleate and palmitate) may also stimulate autophagy (Shen et al., 2012).

Additional nutrient-responsive neuroendocrine mediators, including leptin, adiponectin, ghrelin, myonectin, and others, may affect autophagic responses, establishing a complex network of autophagy-stimulatory and autophagy-inhibitory signals. Prominent autophagy inducers including starvation, physical exercise, rapamycin, resveratrol, and spermidine (a natural polyamine) cause a reduction in circulating leptin levels (He et al., 2012). In the case of exercise, this effect is lost in mice expressing a variant of B cell CLL/lymphoma 2 (BCL2) that blocks stimulus-induced (but not baseline) autophagy (He et al., 2012). Conversely, the reduction of circulating leptin caused by fasting occurs normally in *Atg7*^{-/-} mice (Karsli-Uzunbas et al., 2014). The knockout of adiponectin (*ADIPOQ*), encoding a hormone secreted by the adipose tissue (and placenta), reportedly inhibits autophagy in the myocardium while aggravating diet-induced obesity and the consequent cardiac dysfunction, a series of effects that can be prevented by the administration of rapamycin (Guo et al., 2013b). Ghrelin, also known as the “hunger hormone,” is produced by the gastrointestinal tract when the stomach empties. In vitro, ghrelin mediates both autophagy-activating and autophagy-inhibitory functions, reflecting its ability to inhibit proteasomal protein degradation and AMPK, respectively (Bonfili et al., 2013; Wang et al., 2014). It remains to be determined which of these functions predominate in vivo in physiological versus pathological scenarios. Finally, myonectin (a skeletal muscle-derived hormone encoded by *FAM132B*) suppresses hepatic autophagy upon the stimulation of AKT1/mTORC1 signaling (Seldin et al., 2013). Taken together, these observations indicate that several neuroendocrine mediators regulate autophagy.

Interestingly, autophagy is also coupled to unconventional secretory pathways, mediating the release of a series of soluble

molecules, including diazepam-binding inhibitor (DBI, best known as ACB), interleukin-1 β , and interleukin-18 (Zhang and Schekman, 2013). Thus, autophagy is regulated by several neuroendocrine circuits that sense systemic nutrient availability at the same time that it affects the release of various mediators, including hormones, neurotransmitters, and cytokines. The intricacies of this regulatory network are not yet fully understood and require in-depth exploration.

Despite the limitation of in vitro studies, numerous examples suggest that energy sensors with autophagy-modulatory properties in vitro are also required for autophagy induction by caloric restriction or fasting in vivo. A ketogenic diet (i.e., a high-fat, low-carbohydrate, and low-protein diet supplemented with ketogenic essential amino acids) inhibits mTORC1 in vivo and stimulates autophagy (Xu et al., 2013). Moreover, AMPK is indispensable for myocardial adaptation to caloric restriction in mice (Chen et al., 2013), and sirtuins are required for the autophagy-dependent beneficial effects of nutrient deprivation in nonmammalian model organisms (Morselli et al., 2011) and perhaps in mice (Mercken et al., 2014).

Interestingly, mice that are starved for 24 hr exhibit a significant decrease in cytosolic acetyl-CoA in skeletal and cardiac muscles (but not in the brain) (Mariño et al., 2014), suggesting that the overall nutrient status may affect the abundance of specific intracellular metabolites. Similarly, acetyl-CoA levels drop in the livers of mice experiencing prolonged periods of caloric restriction (Hebert et al., 2013). It is not yet known whether this effect stems directly from decreased nutrient availability or rather reflects a drop in AKT1-dependent activation of ACLY as a consequence of limited INS and IGF1 signaling (Lee et al., 2014).

It appears that (some of the) nutrient sensors originally identified in vitro contribute to the regulation of autophagic responses in vivo. However, we can anticipate that future studies will unravel the major impact of nervous, endocrine, and paracrine signals in the control of autophagy at the whole-body level, contributing to the sophisticated homeostatic regulation that renders the organism adaptable to changes in the quantity and quality of nutrient supplies.

Metabolic Consequences of Autophagy In Individual Cells

Most studies on the metabolic consequences of autophagy compare wild-type cells with cells in which genes encoding essential components of the autophagic machinery have been deleted by homologous recombination or have had their products depleted by RNA interference. Though this approach may provide reliable results in short-term experiments (performed within a few days after knockout or knockdown), it is likely to generate misleading information in long-term settings. Indeed, autophagic defects cause the accumulation of malfunctioning mitochondria and redox-active protein aggregates that, in the long term, have widespread metabolic consequences, including a reduction in mitochondrial ATP synthesis and an increased generation of genotoxic ROS.

In the presence of an intact p53 system, autophagy-deficient cells exhibit impaired proliferation in vivo (Rosenfeldt et al.,

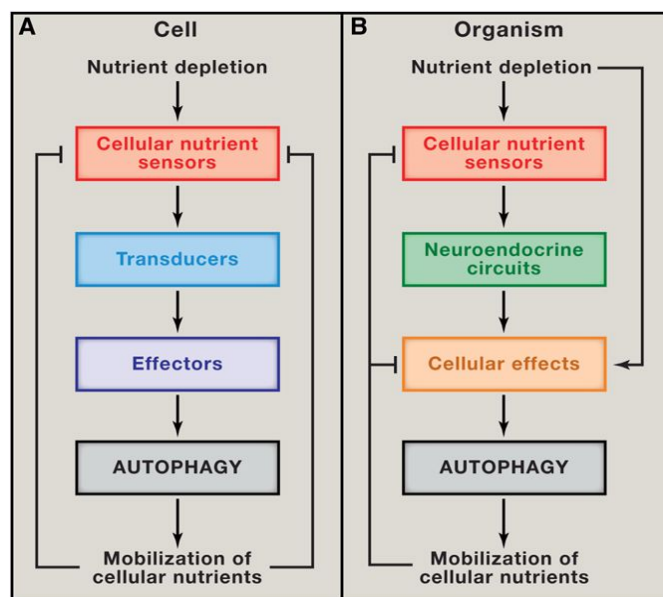


Figure 4. Regulation of Cellular and Organismal Autophagic Responses

(A) When the availability of nutrients in the extracellular fluids drops below a threshold level, sensors expressed by most (if not all) cells are activated and dispatch an autophagy-stimulating stimulus via one or more signal transducer(s). The consequent mobilization of intracellular stores restores, at least to some extent, nutrient availability, resulting in the suppression of autophagy-promoting signals.

(B) Besides mediating direct pro-autophagic effects on virtually all cells, drops in the circulating levels of several nutrients stimulate specific cell types to release neuroendocrine mediators that induce autophagy. These mediators generally trigger autophagic responses in the liver, adipose tissue, or skeletal muscle, resulting in the restoration of systemic nutrient availability and hence in the feedback inhibition of both central and cellular autophagic responses.

2013), perhaps linked to reduced glucose uptake and glycolytic flux (Lock et al., 2011). Moreover, defects in the autophagic machinery result in a marked dependency of Brat^{V600E}-driven lung carcinoma cells on glutamine (Strohecker et al., 2013), whereas they lower mitochondrial oxygen consumption and the levels of Krebs cycle intermediates in the context of KRAS-driven carcinogenesis (Guo et al., 2011).

Autophagy most often delays the transition between a reversible alteration of metabolic homeostasis and the generation of signals that irreversibly commit the cell to death (Galluzzi et al., 2014). Besides counteracting the depletion of energy-rich substrates, which is per se potentially lethal, autophagy limits the accumulation of permeabilized mitochondria, the organelles that regulate the intrinsic pathway of apoptosis and contribute to several instances of regulated necrosis (Galluzzi et al., 2012).

One aspect of autophagy regulation that requires further scrutiny is feedback inhibition. Indeed, it is reasonable to expect that amino acids and other energy-rich metabolites resulting from autophagy-dependent catabolic reactions act on intracellular nutrient sensors to inhibit autophagic responses (Figure 4). By reducing the autophagic flux (irrespective of the conditions that increased it), such a phenomenon may contribute to the shut-down of adaptive responses to stress that marks the recovery of homeostasis.

In Whole Organisms

One problem with the interpretation of data from mouse models bearing tissue-specific knockouts of genes encoding essential components of the autophagic machinery is that autophagy is a key process and its complete suppression invariably entails major metabolic and nonmetabolic alterations. Therefore, studies of the ablation of genes such as *Atg5* and *Atg7* in specific cell types, including subpopulations of hypothalamic neurons, pancreatic β cells, adipocytes, hepatocytes, or myocytes (Kim and Lee, 2014), may provide deeper insights into the impact of degenerative processes affecting such cells rather than the physiological contribution of autophagy to metabolic control.

Recently, the effects of the conditional deletion of *Atg7* at the whole-body level have been explored by expressing a ubiquitous transgene coding for an inducible variant of the Cre recombinase in *Atg7^{lox/lox}* mice. In this setting, tamoxifen administration in the drinking water results in the systemic excision of both floxed *Atg7* alleles (*Atg7^{Δ/Δ}* genotype) (Karsli-Uzunbas et al., 2014). Disabling autophagy in 8- to 10-week-old mice promotes an acute immunodeficiency syndrome that increases the susceptibility of mice to lethal staphylococcal infections (10% mortality 2 weeks after the administration of tamoxifen), and degenerative processes affecting all organs (notably the brain) account for the demise of all animals 6–15 weeks postknockout (Karsli-Uzunbas et al., 2014). Of note, two of the early consequences of *Atg7* deletion are the absence of liver glycogen and the replacement of white with brown adipose tissue. At this stage, i.e., 10 days after knockout, *Atg7^{Δ/Δ}* mice are very sensitive to a short period (24 hr) of starvation. In particular, they exhibit profound hypoglycemia, fail to mobilize fatty acids, and succumb to starvation while manifesting massive hepatic damage (as indicated by a surge in the circulating levels of hepatic enzymes) and severe muscle wasting, as well as DNA damage responses in hepatocytes and myocytes. In this setting, glucose supplementation is sufficient to postpone death and muscle wasting, supporting the notion that *Atg7^{Δ/Δ}* mice truly die from hypoglycemia in response to short periods of starvation (Karsli-Uzunbas et al., 2014). Taken together, these results suggest that the severe consequences of a complete, irreversible inhibition of the autophagic machinery prevent metabolic adaptations to fasting.

The metabolic consequences of partial autophagy inhibition have also been explored. A BCL2 variant in which three amino acids (i.e., T69, S70, and S84) have been substituted with alanine residues (referred to as BCL2^{AAA}) cannot be phosphorylated by JNK1. These mutations prevent the phosphorylation-dependent breakdown of BCL2/BECN1 complexes, hence allowing for baseline, but not stimulus-induced, autophagy (He et al., 2012). Thus, mice expressing BCL2^{AAA} at the whole-body level display lower degrees of exercise-induced autophagy in skeletal muscles than their wild-type counterparts. This correlates with reduced physical endurance and a decreased sensitivity of skeletal muscles to exercise-induced insulin. Moreover, although a high-fat diet causes obesity and type 2 diabetes in both control and BCL2^{AAA}-expressing mice, the beneficial effects of exercise training on diabetes are lost in the latter. Very similar results were obtained in *Becn1^{+/-}* mice (He et al., 2012). These findings suggest that autophagy in skeletal muscles may contribute to the systemic beneficial effects of exercise.

Becn1^{+/-} mice of 16–24 months of age also exhibit increased lipid accumulation in the liver as compared to their age-matched wild-type counterparts (Amir and Czaja, 2011). Moreover, in a mosaic *Atg5* knockout model (in which only a fraction of tissues, including the liver, exhibit autophagic defects), hepatocytes became highly loaded with lipid droplets by 19 months of age (Takamura et al., 2011), supporting the idea that autophagy may counteract steatosis in a cell-autonomous fashion.

The overexpression of essential components of the autophagic machinery may yield useful information regarding the contribution of autophagy to whole-body physiology, provided that such a manipulation is not deleterious for specific cell types. For example, the transgenic expression of *Atg5* under the control of a universal, moderate promoter extends the lifespan of mice by 17.2% (Pyo et al., 2013). This suggests that promoting autophagy to some extent is nontoxic. Importantly, *Atg5*-overexpressing mice not only exhibit a lean phenotype, reduced hepatic levels of triglycerides, increased glucose clearance, and insulin sensitivity as compared to their wild-type counterparts, but are also protected against diet-induced obesity and insulin resistance (Pyo et al., 2013). These findings support a general anti-obesity and anti-diabetes role of autophagy.

Nutritional interventions and other (pharmacological or genetic) inducers of autophagy may also be used to manipulate metabolism *in vivo*. Ample evidence indicates that caloric restriction or specific nutritional manipulations (e.g., methionine restriction, polyamine supplementation) can increase the longevity of nonmammalian organisms (such as yeast, nematodes, and flies) in an autophagy-dependent manner. Similarly, rapamycin and several other experimental inducers of autophagy (such as resveratrol and spermidine) extend the lifespan of nonmammalian species in an autophagy-dependent fashion (Rubinshtein et al., 2011). Rapamycin also increases the longevity of laboratory mice and reverses age-related cardiac dysfunction, even when administered late in life (Harrison et al., 2009). However, it has not yet been determined whether autophagy is required for the lifespan-extending effects of rapamycin on mice. Moreover, rapamycin may promote type 2 diabetes, probably due to inhibition of mTORC2 (Lamming et al., 2012).

Future studies must determine which, if any, dosing schedule might reduce the immunosuppressive and pro-diabetic side effects of rapamycin or whether rapamycin should be replaced by more specific mTORC1 inhibitors (Li et al., 2014), other drugs that mimic the effects of caloric restriction (i.e., that reduce acetyl-CoA levels, such as ACLY inhibitors; stimulate sirtuins, such as nicotinamide or resveratrol; and/or inhibit acetyltransferases, such as spermidine) (Madeo et al., 2014), or agents that specifically activate the autophagic machinery (such as a cell-permeable peptide that derepresses BECN1) (Shoji-Kawata et al., 2013). Indeed, although accumulating evidence indicates that nicotinamide and resveratrol have profound anti-obesity and anti-diabetes effects (Cantó et al., 2012; Lagouge et al., 2006), it remains elusive to what extent such activity is mediated by autophagy.

Both rapamycin and nicotinamide riboside can be used to treat experimental mitochondrial pathologies. In particular, rapamycin improves the clinical course of mice lacking a mitochondrial

respiratory chain subunit (NADH dehydrogenase [ubiquinone] Fe-S protein 4, 18 kDa, *Ndufs4*), which develop brain lesions similar to those associated with the Leigh syndrome (Johnson et al., 2013). Similarly, nicotinamide riboside causes a significant improvement in the symptoms of mice carrying a dominant mutation in the gene coding for progressive external ophthalmoplegia 1 (*Peo1*, a mitochondrial replicative helicase), which results in progressive mitochondrial myopathy upon the accumulation of mutated mtDNA (Khan et al., 2014). Moreover, nicotinamide riboside improves the mitochondrial defects and intolerance to physical exercise of mice expressing a pathologic variant of SCO cytochrome oxidase-deficient homolog 2 (*Sco2*) in the absence of endogenous *Sco2* alleles, which develop a mitochondrial disease model characterized by impaired biogenesis of cytochrome c oxidase (Cerutti et al., 2014). Because autophagy plays a major role in mitochondrial quality control, a chronic increase in autophagy could contribute to the beneficial effects of these compounds on diseases with limited therapeutic options. This possibility remains to be addressed.

In summary, we are currently witnessing the development of new pharmacological and genetic methods to manipulate (induce or suppress) autophagy, which should facilitate the exploration of this key pathway in physiological scenarios. It will be particularly interesting to examine autophagy by techniques that permit its partial and transient modulation, either in the entire organism or in defined organs. Beyond mechanistic insights, such an approach will yield information on the therapeutic utility and potential long-term side effects of autophagy-modulating measures.

Conclusions and Perspectives

Undoubtedly, metabolism regulates autophagy, and autophagy has a profound impact on metabolism. As a major manifestation of this tight interrelationship, the autophagy-dependent mobilization of cellular and organismal reserves triggers negative feedback circuitries that inhibit autophagy at both the single-cell and systemic level (Figure 4). Autophagy is a complex process that requires a major degree of coordination among distinct molecular systems, ensuring that the initial sequestration of the autophagic cargo in autophagosomes leads to lysosomal degradation. Indeed, the induction of autophagy by heterogeneous interventions provokes a relatively homogeneous response characterized by the activation of specific kinases (AMPK, IKK, JNK1, TAK1, ULK1, VPS34), the inhibition of others (such as mTORC1), protein deacetylation reactions (at least in part ensuring the activation of SIRT1 and/or the inhibition of EP300), and the reversal of inhibitory interactions such as those between BECN1 and Bcl-2 family members.

Such a tight coordination may be achieved by several mechanisms, including (but not limited to): (1) mutually stimulatory pro-autophagic interactions among nutrient sensors (e.g., AMPK and SIRT1, AMPK and mTORC1), (2) the direct activation of several pro-autophagic factors by nutrient sensors (e.g., ULK1 and TFEB, which are regulated by mTORC1), (3) positive interactions among essential molecules involved in distinct steps of the autophagic process (e.g., the phosphorylation of BECN1 and AMBRA1 by ULK1; the phosphorylation of AMBRA1 by

mTORC1); and (4) synchronized rearrangements of key factors in functionally distinct multiprotein complexes. This may explain why distinct primary signals can stimulate stereotyped changes in several supramolecular complexes involved in the regulation of autophagy.

An emerging theme is that autophagy responds to the depletion of a panel of nutrients by mobilizing intracellular reserves. Still, it is not clear yet whether distinct types of nutrient deficiencies may cause a highly specific and graduated autophagic response resembling the one triggered by iron deficiency. Thus, one might speculate that fluctuations in the abundance of specific nutrients might stimulate an autophagic response (in terms of autophagic cargo and cell types involved) that differs from the one induced by indiscriminate caloric restriction. Irrespective of these and other unknowns, autophagy exerts major homeostatic control on both cellular and organismal metabolism. Thus, we anticipate that pathological alterations of autophagy and their therapeutic correction will occupy a central stage in future clinical practice.

SUPPLEMENTAL INFORMATION

Supplemental Information includes one table and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2014.11.006>.

ACKNOWLEDGMENTS

We are indebted to Erika Vacchelli (Centre de Recherche des Cordeliers) for help with table preparation. G.K. is supported by the Ligue contre le Cancer (équipe labélisée); Agence Nationale de la Recherche (ANR); Association pour la recherche sur le cancer (ARC); Cancéropôle Ile-de-France; Institut National du Cancer (INCa); Fondation Bettencourt-Schueller; Fondation de France; Fondation pour la Recherche Médicale (FRM); the European Commission (ArtForce); the European Research Council (ERC); the E-Rare-2 project, the LabEx Immuno-Oncology; the SIRIC Stratified Oncology Cell DNA Repair and Tumor Immune Elimination (SOCRATE); the SIRIC Cancer Research and Personalized Medicine (CARPEM); and the Paris Alliance of Cancer Research Institutes (PACRI). B.L. is supported by NIH grants RO1 CA109618 and U19 AI109725 and a Cancer Prevention Research Institute of Texas grant RP120718-PI.

REFERENCES

- Amir, M., and Czaja, M.J. (2011). Autophagy in nonalcoholic steatohepatitis. *Expert Rev. Gastroenterol. Hepatol.* 5, 159–166.
- Atkinson, D.E., and Walton, G.M. (1967). Adenosine triphosphate conservation in metabolic regulation. Rat liver citrate cleavage enzyme. *J. Biol. Chem.* 242, 3239–3241.
- B'chir, W., Maurin, A.C., Carraro, V., Averous, J., Jousse, C., Muranishi, Y., Parry, L., Stepien, G., Fafournoux, P., and Bruhat, A. (2013). The eIF2 α /ATF4 pathway is essential for stress-induced autophagy gene expression. *Nucleic Acids Res.* 41, 7683–7699.
- Bonfili, L., Cuccioloni, M., Cecarini, V., Mozzicafreddo, M., Palermo, F.A., Cocci, P., Angeletti, M., and Eleuteri, A.M. (2013). Ghrelin induces apoptosis in colon adenocarcinoma cells via proteasome inhibition and autophagy induction. *Apoptosis* 18, 1188–1200.
- Cantó, C., Gerhart-Hines, Z., Feige, J.N., Lagouge, M., Noriega, L., Milne, J.C., Elliott, P.J., Puigserver, P., and Auwerx, J. (2009). AMPK regulates energy expenditure by modulating NAD⁺ metabolism and SIRT1 activity. *Nature* 458, 1056–1060.
- Cantó, C., Houtkooper, R.H., Pirinen, E., Youn, D.Y., Oosterveer, M.H., Cen, Y., Fernandez-Marcos, P.J., Yamamoto, H., Andreux, P.A., Cettour-Rose, P., et al. (2012). The NAD(+) precursor nicotinamide riboside enhances oxidative metabolism and protects against high-fat diet-induced obesity. *Cell Metab.* 15, 838–847.
- Cerutti, R., Pirinen, E., Lamperti, C., Marchet, S., Sauve, A.A., Li, W., Leoni, V., Schon, E.A., Dantzer, F., Auwerx, J., et al. (2014). NAD(+) dependent activation of Sirt1 corrects the phenotype in a mouse model of mitochondrial disease. *Cell Metab.* 19, 1042–1049.
- Chalkiadaki, A., and Guarente, L. (2012). Sirtuins mediate mammalian metabolic responses to nutrient availability. *Nat. Rev. Endocrinol.* 8, 287–296.
- Chen, K., Kobayashi, S., Xu, X., Viollet, B., and Liang, Q. (2013). AMP activated protein kinase is indispensable for myocardial adaptation to caloric restriction in mice. *PLoS ONE* 8, e59682.
- Cheng, H.L., Mostoslavsky, R., Saito, S., Manis, J.P., Gu, Y., Patel, P., Bronson, R., Appella, E., Alt, F.W., and Chua, K.F. (2003). Developmental defects and p53 hyperacetylation in Sir2 homolog (SIRT1)-deficient mice. *Proc. Natl. Acad. Sci. USA* 100, 10794–10799.
- Cheng, C.W., Adams, G.B., Perin, L., Wei, M., Zhou, X., Lam, B.S., Da Sacco, S., Mirisola, M., Quinn, D.I., Dorff, T.B., et al. (2014). Prolonged fasting reduces IGF-1/PKA to promote hematopoietic-stem-cell-based regeneration and reverse immunosuppression. *Cell Stem Cell* 14, 810–823.
- Cheong, H., Lindsten, T., Wu, J., Lu, C., and Thompson, C.B. (2011). Ammonia-induced autophagy is independent of ULK1/ULK2 kinases. *Proc. Natl. Acad. Sci. USA* 108, 11121–11126.
- Chin, R.M., Fu, X., Pai, M.Y., Vergnes, L., Hwang, H., Deng, G., Diep, S., Lomenick, B., Meli, V.S., Monsalve, G.C., et al. (2014). The metabolite α -ketoglutarate extends lifespan by inhibiting ATP synthase and TOR. *Nature* 510, 397–401.
- Choi, A.M., Ryter, S.W., and Levine, B. (2013). Autophagy in human health and disease. *N. Engl. J. Med.* 368, 651–662.
- Comb, W.C., Hutti, J.E., Cogswell, P., Cantley, L.C., and Baldwin, A.S. (2012). p85 α SH2 domain phosphorylation by IKK promotes feedback inhibition of PI3K and Akt in response to cellular starvation. *Mol. Cell* 45, 719–730.
- Criollo, A., Niso-Santano, M., Malik, S.A., Michaud, M., Morselli, E., Mariño, G., Lachkar, S., Arkhipenko, A.V., Harper, F., Pierron, G., et al. (2011). Inhibition of autophagy by TAB2 and TAB3. *EMBO J.* 30, 4908–4920.
- Dou, Z., Pan, J.A., Dbouk, H.A., Ballou, L.M., DeLeon, J.L., Fan, Y., Chen, J.S., Liang, Z., Li, G., Backer, J.M., et al. (2013). Class IA PI3K p110 β subunit promotes autophagy through Rab5 small GTPase in response to growth factor limitation. *Mol. Cell* 50, 29–42.
- Dupont, N., Chauhan, S., Arko-Mensah, J., Castillo, E.F., Masedunskas, A., Weigert, R., Robenek, H., Proikas-Cezanne, T., and Deretic, V. (2014). Neutral lipid stores and lipase PNPLA5 contribute to autophagosome biogenesis. *Curr. Biol.* 24, 609–620.
- Durán, R.V., Oppliger, W., Robitaille, A.M., Heiserich, L., Skendaj, R., Gottlieb, E., and Hall, M.N. (2012). Glutaminolysis activates Rag-mTORC1 signaling. *Mol. Cell* 47, 349–358.
- Durán, R.V., MacKenzie, E.D., Boulahbel, H., Frezza, C., Heiserich, L., Tardito, S., Bussolati, O., Rocha, S., Hall, M.N., and Gottlieb, E. (2013). HIF-independent role of prolyl hydroxylases in the cellular response to amino acids. *Oncogene* 32, 4549–4556.
- Efeyan, A., Zoncu, R., Chang, S., Gumper, I., Snitkin, H., Wolfson, R.L., Kirak, O., Sabatini, D.D., and Sabatini, D.M. (2013). Regulation of mTORC1 by the Rag GTPases is necessary for neonatal autophagy and survival. *Nature* 493, 679–683.
- Eisenberg, T., Knauer, H., Schauer, A., Büttner, S., Ruckenstein, C., Carmona-Gutierrez, D., Ring, J., Schroeder, S., Magnes, C., Antonacci, L., et al. (2009). Induction of autophagy by spermidine promotes longevity. *Nat. Cell Biol.* 11, 1305–1314.
- Eisenberg, T., Schroeder, S., Andryushkova, A., Pendl, T., Küttner, V., Bhukel, A., Mariño, G., Pietrocola, F., Harger, A., Zimmermann, A., et al. (2014). Nucleocytosolic depletion of the energy metabolite acetyl-coenzyme A stimulates autophagy and prolongs lifespan. *Cell Metab.* 19, 431–444.

- Eng, C.H., Yu, K., Lucas, J., White, E., and Abraham, R.T. (2010). Ammonia derived from glutaminolysis is a diffusible regulator of autophagy. *Sci. Signal.* **3**, ra31.
- Galluzzi, L., Kepp, O., and Kroemer, G. (2012). Mitochondria: master regulators of danger signalling. *Nat. Rev. Mol. Cell Biol.* **13**, 780–788.
- Galluzzi, L., Kepp, O., Vander Heiden, M.G., and Kroemer, G. (2013). Metabolic targets for cancer therapy. *Nat. Rev. Drug Discov.* **12**, 829–846.
- Galluzzi, L., Bravo-San Pedro, J.M., Vitale, I., Aaronson, S.A., Abrams, J.M., Adam, D., Alnemri, E.S., Altucci, L., Andrews, D., Annicchiarico-Petruzzelli, M., et al. (2014). Essential versus accessory aspects of cell death: recommendations of the NCCD 2015. *Cell Death Differ.*, in press.
- Gibson, B.A., and Kraus, W.L. (2012). New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs. *Nat. Rev. Mol. Cell Biol.* **13**, 411–424.
- Green, D.R., Galluzzi, L., and Kroemer, G. (2011). Mitochondria and the autophagy-inflammation-cell death axis in organismal aging. *Science* **333**, 1109–1112.
- Guo, J.Y., Chen, H.Y., Mathew, R., Fan, J., Strohecker, A.M., Karsli-Uzunbas, G., Kamphorst, J.J., Chen, G., Lemons, J.M., Karantza, V., et al. (2011). Activated Ras requires autophagy to maintain oxidative metabolism and tumorigenesis. *Genes Dev.* **25**, 460–470.
- Guo, J.Y., Xia, B., and White, E. (2013a). Autophagy-mediated tumor promotion. *Cell* **155**, 1216–1219.
- Guo, R., Zhang, Y., Turdi, S., and Ren, J. (2013b). Adiponectin knockout accentuates high fat diet-induced obesity and cardiac dysfunction: role of autophagy. *Biochim. Biophys. Acta* **1832**, 1136–1148.
- Han, J.M., Jeong, S.J., Park, M.C., Kim, G., Kwon, N.H., Kim, H.K., Ha, S.H., Ryu, S.H., and Kim, S. (2012). Leucyl-tRNA synthetase is an intracellular leucine sensor for the mTORC1-signaling pathway. *Cell* **149**, 410–424.
- Harder, L.M., Bunkenborg, J., and Andersen, J.S. (2014). Inducing autophagy: a comparative phosphoproteomic study of the cellular response to ammonia and rapamycin. *Autophagy* **10**, 339–355.
- Hardie, D.G., Ross, F.A., and Hawley, S.A. (2012). AMPK: a nutrient and energy sensor that maintains energy homeostasis. *Nat. Rev. Mol. Cell Biol.* **13**, 251–262.
- Harrison, D.E., Strong, R., Sharp, Z.D., Nelson, J.F., Astle, C.M., Flurkey, K., Nadon, N.L., Wilkinson, J.E., Frenkel, K., Carter, C.S., et al. (2009). Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. *Nature* **460**, 392–395.
- Hawley, S.A., Fullerton, M.D., Ross, F.A., Schertzer, J.D., Chevztzoff, C., Walker, K.J., Pegg, M.W., Zibrova, D., Green, K.A., Mustard, K.J., et al. (2012). The ancient drug salicylate directly activates AMP-activated protein kinase. *Science* **336**, 918–922.
- He, C., Bassik, M.C., Moresi, V., Sun, K., Wei, Y., Zou, Z., An, Z., Loh, J., Fisher, J., Sun, Q., et al. (2012). Exercise-induced BCL2-regulated autophagy is required for muscle glucose homeostasis. *Nature* **481**, 511–515.
- Hebert, A.S., Dittenhafer-Reed, K.E., Yu, W., Bailey, D.J., Selen, E.S., Boersma, M.D., Carson, J.J., Tonelli, M., Balloon, A.J., Higbee, A.J., et al. (2013). Calorie restriction and SIRT3 trigger global reprogramming of the mitochondrial protein acetylome. *Mol. Cell* **49**, 186–199.
- Houtkooper, R.H., Pirinen, E., and Auwerx, J. (2012). Sirtuins as regulators of metabolism and healthspan. *Nat. Rev. Mol. Cell Biol.* **13**, 225–238.
- Inokuchi-Shimizu, S., Park, E.J., Roh, Y.S., Yang, L., Zhang, B., Song, J., Liang, S., Pimienta, M., Taniguchi, K., Wu, X., et al. (2014). TAK1-mediated autophagy and fatty acid oxidation prevent hepatosteatosis and tumorigenesis. *J. Clin. Invest.* **124**, 3566–3578.
- Johnson, S.C., Yanos, M.E., Kayser, E.B., Quintana, A., Sangesland, M., Castanza, A., Uhde, L., Hui, J., Wall, V.Z., Gagnidze, A., et al. (2013). mTOR inhibition alleviates mitochondrial disease in a mouse model of Leigh syndrome. *Science* **342**, 1524–1528.
- Karsli-Uzunbas, G., Guo, J.Y., Price, S., Teng, X., Laddha, S.V., Khor, S., Kalaany, N.Y., Jacks, T., Chan, C.S., Rabinowitz, J.D., and White, E. (2014). Autophagy is required for glucose homeostasis and lung tumor maintenance. *Cancer Discov.* **4**, 914–927.
- Khalidoun, S.A., Emond-Boisjoly, M.A., Chateau, D., Carrière, V., Lacasa, M., Rousset, M., Demignot, S., and Morel, E. (2014). Autophagosomes contribute to intracellular lipid distribution in enterocytes. *Mol. Biol. Cell* **25**, 118–132.
- Khan, N.A., Auranen, M., Paetau, I., Pirinen, E., Euro, L., Forsström, S., Pasila, L., Velagapudi, V., Carroll, C.J., Auwerx, J., and Suomalainen, A. (2014). Effective treatment of mitochondrial myopathy by nicotinamide riboside, a vitamin B3. *EMBO Mol. Med.* **6**, 721–731.
- Kim, K.H., and Lee, M.S. (2014). Autophagy—a key player in cellular and body metabolism. *Nat. Rev. Endocrinol.* **10**, 322–337.
- Kim, J., Kundu, M., Viollet, B., and Guan, K.L. (2011). AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1. *Nat. Cell Biol.* **13**, 132–141.
- Kim, J., Kim, Y.C., Fang, C., Russell, R.C., Kim, J.H., Fan, W., Liu, R., Zhong, Q., and Guan, K.L. (2013). Differential regulation of distinct Vps34 complexes by AMPK in nutrient stress and autophagy. *Cell* **152**, 290–303.
- Kondomerkos, D.J., Kalamidas, S.A., Kotoulas, O.B., and Hann, A.C. (2005). Glycogen autophagy in the liver and heart of newborn rats. The effects of glucagon, adrenalin or rapamycin. *Histol. Histopathol.* **20**, 689–696.
- Kroemer, G., Mariño, G., and Levine, B. (2010). Autophagy and the integrated stress response. *Mol. Cell* **40**, 280–293.
- Kuma, A., Hatano, M., Matsui, M., Yamamoto, A., Nakaya, H., Yoshimori, T., Ohsumi, Y., Tokuhi, T., and Mizushima, N. (2004). The role of autophagy during the early neonatal starvation period. *Nature* **432**, 1032–1036.
- Lagouge, M., Argmann, C., Gerhart-Hines, Z., Meziane, H., Lerin, C., Daussin, F., Messadeq, N., Milne, J., Lambert, P., Elliott, P., et al. (2006). Resveratrol improves mitochondrial function and protects against metabolic disease by activating SIRT1 and PGC-1 α . *Cell* **127**, 1109–1122.
- Lamming, D.W., Ye, L., Katajisto, P., Goncalves, M.D., Saitoh, M., Stevens, D.M., Davis, J.G., Salmon, A.B., Richardson, A., Ahima, R.S., et al. (2012). Rapamycin-induced insulin resistance is mediated by mTORC2 loss and uncoupled from longevity. *Science* **335**, 1638–1643.
- Lanna, A., Henson, S.M., Escors, D., and Akbar, A.N. (2014). The kinase p38 activated by the metabolic regulator AMPK and scaffold TAB1 drives the senescence of human T cells. *Nat. Immunol.* **15**, 965–972.
- Lapierre, L.R., De Magalhães Filho, C.D., McQuary, P.R., Chu, C.C., Visvikis, O., Chang, J.T., Gelino, S., Ong, B., Davis, A.E., et al. (2013). The TFEB orthologue HLH-30 regulates autophagy and modulates longevity in *Caenorhabditis elegans*. *Nat. Commun.* **4**, 2267.
- Lee, I.H., and Finkel, T. (2009). Regulation of autophagy by the p300 acetyltransferase. *J. Biol. Chem.* **284**, 6322–6328.
- Lee, I.H., Cao, L., Mostoslavsky, R., Lombard, D.B., Liu, J., Bruns, N.E., Tsokos, M., Alt, F.W., and Finkel, T. (2008). A role for the NAD-dependent deacetylase Sirt1 in the regulation of autophagy. *Proc. Natl. Acad. Sci. USA* **105**, 3374–3379.
- Lee, J.V., Carrer, A., Shah, S., Snyder, N.W., Wei, S., Venneti, S., Worth, A.J., Yuan, Z.F., Lim, H.W., Liu, S., et al. (2014). Akt-dependent metabolic reprogramming regulates tumor cell histone acetylation. *Cell Metab.* **20**, 306–319.
- Li, N., Wu, X., Holzer, R.G., Lee, J.H., Todoric, J., Park, E.J., Ogata, H., Gukovskaya, A.S., Gukovsky, I., Pizzo, D.P., et al. (2013). Loss of acinar cell IKK α triggers spontaneous pancreatitis in mice. *J. Clin. Invest.* **123**, 2231–2243.
- Li, J., Kim, S.G., and Blenis, J. (2014). Rapamycin: one drug, many effects. *Cell Metab.* **19**, 373–379.
- Lizaso, A., Tan, K.T., and Lee, Y.H. (2013). β -adrenergic receptor-stimulated lipolysis requires the RAB7-mediated autolysosomal lipid degradation. *Autophagy* **9**, 1228–1243.
- Lock, R., Roy, S., Kenific, C.M., Su, J.S., Salas, E., Ronen, S.M., and Debnath, J. (2011). Autophagy facilitates glycolysis during Ras-mediated oncogenic transformation. *Mol. Biol. Cell* **22**, 165–178.

- Mackeh, R., Lorin, S., Ratier, A., Mejdoubi-Charef, N., Baillet, A., Bruneel, A., Hamai, A., Codogno, P., Poüs, C., and Perdiz, D. (2014). Reactive oxygen species, AMP-activated protein kinase, and the transcription cofactor p300 regulate α -tubulin acetyltransferase-1 (α TAT-1/MEC-17)-dependent microtubule hyperacetylation during cell stress. *J. Biol. Chem.* 289, 11816–11828.
- Madeo, F., Pietrocola, F., Eisenberg, T., and Kroemer, G. (2014). Caloric restriction mimetics: towards a molecular definition. *Nat. Rev. Drug Discov.* 13, 727–740.
- Mader, B.J., Pivtoraiko, V.N., Flippo, H.M., Klocke, B.J., Roth, K.A., Mangieri, L.R., and Shacka, J.J. (2012). Rotenone inhibits autophagic flux prior to inducing cell death. *ACS Chem. Neurosci.* 3, 1063–1072.
- Mancias, J.D., Wang, X., Gygi, S.P., Harper, J.W., and Kimmelman, A.C. (2014). Quantitative proteomics identifies NCOA4 as the cargo receptor mediating ferritinophagy. *Nature* 509, 105–109.
- Mariño, G., Pietrocola, F., Eisenberg, T., Kong, Y., Malik, S.A., Andryushkova, A., Schroeder, S., Pendl, T., Harger, A., Niso-Santano, M., et al. (2014). Regulation of autophagy by cytosolic acetyl-coenzyme A. *Mol. Cell* 53, 710–725.
- Mercken, E.M., Hu, J., Krzysik-Walker, S., Wei, M., Li, Y., McBurney, M.W., de Cabo, R., and Longo, V.D. (2014). SIRT1 but not its increased expression is essential for lifespan extension in caloric-restricted mice. *Aging Cell* 13, 193–196.
- Mizushima, N. (2007). Autophagy: process and function. *Genes Dev.* 21, 2861–2873.
- Mizushima, N. (2009). Methods for monitoring autophagy using GFP-LC3 transgenic mice. *Methods Enzymol.* 452, 13–23.
- Mizushima, N., and Komatsu, M. (2011). Autophagy: renovation of cells and tissues. *Cell* 147, 728–741.
- Morselli, E., Mariño, G., Bernetzen, M.V., Eisenberg, T., Megalou, E., Schroeder, S., Cabrera, S., Bénit, P., Rustin, P., Criollo, A., et al. (2011). Spermidine and resveratrol induce autophagy by distinct pathways converging on the acetylproteome. *J. Cell Biol.* 192, 615–629.
- Nazio, F., Strappazzon, F., Antonioli, M., Bielli, P., Cianfanelli, V., Bordin, M., Gretzmeier, C., Dengjel, J., Piacentini, M., Fimia, G.M., and Cecconi, F. (2013). mTOR inhibits autophagy by controlling ULK1 ubiquitylation, self-association and function through AMBRA1 and TRAF6. *Nat. Cell Biol.* 15, 406–416.
- Ogasawara, Y., Itakura, E., Kono, N., Mizushima, N., Arai, H., Nara, A., Mizukami, T., and Yamamoto, A. (2014). Stearoyl-CoA desaturase 1 activity is required for autophagosome formation. *J. Biol. Chem.* 289, 23938–23950.
- Pyo, J.O., Yoo, S.M., Ahn, H.H., Nah, J., Hong, S.H., Kam, T.I., Jung, S., and Jung, Y.K. (2013). Overexpression of Atg5 in mice activates autophagy and extends lifespan. *Nat. Commun.* 4, 2300.
- Riehle, C., Wende, A.R., Sena, S., Pires, K.M., Pereira, R.O., Zhu, Y., Bugger, H., Frank, D., Bevins, J., Chen, D., et al. (2013). Insulin receptor substrate signaling suppresses neonatal autophagy in the heart. *J. Clin. Invest.* 123, 5319–5333.
- Roberts, D.J., Tan-Sah, V.P., Ding, E.Y., Smith, J.M., and Miyamoto, S. (2014). Hexokinase-II positively regulates glucose starvation-induced autophagy through TORC1 inhibition. *Mol. Cell* 53, 521–533.
- Rosenfeldt, M.T., O'Prey, J., Morton, J.P., Nixon, C., MacKay, G., Mrowinska, A., Au, A., Rai, T.S., Zheng, L., Ridgway, R., et al. (2013). p53 status determines the role of autophagy in pancreatic tumour development. *Nature* 504, 296–300.
- Rubinsztein, D.C., Mariño, G., and Kroemer, G. (2011). Autophagy and aging. *Cell* 146, 682–695.
- Sancak, Y., Bar-Peled, L., Zoncu, R., Markhard, A.L., Nada, S., and Sabatini, D.M. (2010). Ragulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. *Cell* 141, 290–303.
- Scherz-Shouval, R., and Elazar, Z. (2011). Regulation of autophagy by ROS: physiology and pathology. *Trends Biochem. Sci.* 36, 30–38.
- Seldin, M.M., Lei, X., Tan, S.Y., Stanson, K.P., Wei, Z., and Wong, G.W. (2013). Skeletal muscle-derived myonectin activates the mammalian target of rapamycin (mTOR) pathway to suppress autophagy in liver. *J. Biol. Chem.* 288, 36073–36082.
- Settembre, C., Fraldi, A., Medina, D.L., and Ballabio, A. (2013). Signals from the lysosome: a control centre for cellular clearance and energy metabolism. *Nat. Rev. Mol. Cell Biol.* 14, 283–296.
- Sharara-Chami, R.I., Zhou, Y., Ebert, S., Pacak, K., Ozcan, U., and Majzoub, J.A. (2012). Epinephrine deficiency results in intact glucose counter-regulation, severe hepatic steatosis and possible defective autophagy in fasting mice. *Int. J. Biochem. Cell Biol.* 44, 905–913.
- Shen, S., Niso-Santano, M., Adjemian, S., Takehara, T., Malik, S.A., Minoux, H., Souquere, S., Mariño, G., Lachkar, S., Senovilla, L., et al. (2012). Cytoplasmic STAT3 represses autophagy by inhibiting PKR activity. *Mol. Cell* 48, 667–680.
- Shimobayashi, M., and Hall, M.N. (2014). Making new contacts: the mTOR network in metabolism and signalling crosstalk. *Nat. Rev. Mol. Cell Biol.* 15, 155–162.
- Shoji-Kawata, S., Sumpter, R., Leveno, M., Campbell, G.R., Zou, Z., Kinch, L., Wilkins, A.D., Sun, Q., Pallau, K., MacDuff, D., et al. (2013). Identification of a candidate therapeutic autophagy-inducing peptide. *Nature* 494, 201–206.
- Silvera, D., Formenti, S.C., and Schneider, R.J. (2010). Translational control in cancer. *Nat. Rev. Cancer* 10, 254–266.
- Singh, R., and Cuervo, A.M. (2011). Autophagy in the cellular energetic balance. *Cell Metab.* 13, 495–504.
- Strohecker, A.M., Guo, J.Y., Karsli-Uzunbas, G., Price, S.M., Chen, G.J., Mathew, R., McMahon, M., and White, E. (2013). Autophagy sustains mitochondrial glutamine metabolism and growth of BrafV600E-driven lung tumors. *Cancer Discov.* 3, 1272–1285.
- Takamura, A., Komatsu, M., Hara, T., Sakamoto, A., Kishi, C., Waguri, S., Eishi, Y., Hino, O., Tanaka, K., and Mizushima, N. (2011). Autophagy-deficient mice develop multiple liver tumors. *Genes Dev.* 25, 795–800.
- Tallóczy, Z., Jiang, W., Virgin, H.W., 4th, Leib, D.A., Scheuner, D., Kaufman, R.J., Eskelinen, E.L., and Levine, B. (2002). Regulation of starvation- and virus-induced autophagy by the eIF2 α kinase signaling pathway. *Proc. Natl. Acad. Sci. USA* 99, 190–195.
- Troncoso, R., Vicencio, J.M., Parra, V., Nemchenko, A., Kawashima, Y., Del Campo, A., Toro, B., Battiprolu, P.K., Aranguiz, P., Chiong, M., et al. (2012). Energy-preserving effects of IGF-1 antagonize starvation-induced cardiac autophagy. *Cardiovasc. Res.* 93, 320–329.
- Wang, R.C., Wei, Y., An, Z., Zou, Z., Xiao, G., Bhagat, G., White, M., Reichelt, J., and Levine, B. (2012). Akt-mediated regulation of autophagy and tumorigenesis through Beclin 1 phosphorylation. *Science* 338, 956–959.
- Wang, X., Wang, X.L., Chen, H.L., Wu, D., Chen, J.X., Wang, X.X., Li, R.L., He, J.H., Mo, L., Cen, X., et al. (2014). Ghrelin inhibits doxorubicin cardiotoxicity by inhibiting excessive autophagy through AMPK and p38-MAPK. *Biochem. Pharmacol.* 88, 334–350.
- Wauson, E.M., Zaganjor, E., Lee, A.Y., Guerra, M.L., Ghosh, A.B., Bookout, A.L., Chambers, C.P., Jivan, A., McGlynn, K., Hutchison, M.R., et al. (2012). The G protein-coupled taste receptor T1R1/T1R3 regulates mTORC1 and autophagy. *Mol. Cell* 47, 851–862.
- Wauson, E.M., Dbouk, H.A., Ghosh, A.B., and Cobb, M.H. (2014). G protein-coupled receptors and the regulation of autophagy. *Trends Endocrinol. Metab.* 25, 274–282.
- Wei, Y., Zou, Z., Becker, N., Anderson, M., Sumpter, R., Xiao, G., Kinch, L., Koduru, P., Christudass, C.S., Veltri, R.W., et al. (2013). EGFR-mediated Beclin 1 phosphorylation in autophagy suppression, tumor progression, and tumor chemoresistance. *Cell* 154, 1269–1284.
- Wieman, H.L., Wofford, J.A., and Rathmell, J.C. (2007). Cytokine stimulation promotes glucose uptake via phosphatidylinositol-3 kinase/Akt regulation of Glut1 activity and trafficking. *Mol. Biol. Cell* 18, 1437–1446.
- Williams-Bey, Y., Boularan, C., Vural, A., Huang, N.N., Hwang, I.Y., Shan-Shi, C., and Kehrl, J.H. (2014). Omega-3 free fatty acids suppress macrophage inflammasome activation by inhibiting NF- κ B activation and enhancing autophagy. *PLoS ONE* 9, e97957.

- Xu, L., Kanasaki, M., He, J., Kitada, M., Nagao, K., Jinzu, H., Noguchi, Y., Maegawa, H., Kanasaki, K., and Koya, D. (2013). Ketogenic essential amino acids replacement diet ameliorated hepatosteatosis with altering autophagy-associated molecules. *Biochim. Biophys. Acta* 1832, 1605–1612.
- Yang, W., Hong, Y.H., Shen, X.Q., Frankowski, C., Camp, H.S., and Leff, T. (2001). Regulation of transcription by AMP-activated protein kinase: phosphorylation of p300 blocks its interaction with nuclear receptors. *J. Biol. Chem.* 276, 38341–38344.
- Yang, L., Li, P., Fu, S., Calay, E.S., and Hotamisligil, G.S. (2010). Defective hepatic autophagy in obesity promotes ER stress and causes insulin resistance. *Cell Metab.* 11, 467–478.
- Ye, J., Kumanova, M., Hart, L.S., Sloane, K., Zhang, H., De Panis, D.N., Bobrovnikova-Marjon, E., Diehl, J.A., Ron, D., and Koumenis, C. (2010). The GCN2-ATF4 pathway is critical for tumour cell survival and proliferation in response to nutrient deprivation. *EMBO J.* 29, 2082–2096.
- Yuan, H.X., Russell, R.C., and Guan, K.L. (2013). Regulation of PIK3C3/VPS34 complexes by MTOR in nutrient stress-induced autophagy. *Autophagy* 9, 1983–1995.
- Zhang, M., and Schekman, R. (2013). Cell biology. Unconventional secretion, unconventional solutions. *Science* 340, 559–561.
- Zoncu, R., Bar-Peled, L., Efeyan, A., Wang, S., Sancak, Y., and Sabatini, D.M. (2011). mTORC1 senses lysosomal amino acids through an inside-out mechanism that requires the vacuolar H(+)-ATPase. *Science* 334, 678–683.

Gut Microbiota Elicits a Protective Immune Response against Malaria Transmission

Bahtiyar Yilmaz,¹ Silvia Portugal,² Tuan M. Tran,² Raffaella Gozzelino,¹ Susana Ramos,¹ Joana Gomes,^{1,3} Ana Regalado,¹ Peter J. Cowan,^{4,5} Anthony J.F. d'Apice,^{4,5} Anita S. Chong,⁶ Ogobara K. Doumbo,⁷ Boubacar Traore,⁷ Peter D. Crompton,² Henrique Silveira,³ and Miguel P. Soares^{1,*}

¹Instituto Gulbenkian de Ciência, Rua da Quinta Grande, 6, 2780-156 Oeiras, Portugal

²Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Twinbrook II, Room 125, 12441 Parklawn Drive, Rockville, MD 20852-8180, USA

³Centro de Malaria e Outras Doenças Tropicais, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Rua da Junqueira, 100, 1349-008 Lisboa, Portugal

⁴Immunology Research Centre, St. Vincent's Hospital, Fitzroy, Melbourne, VIC 3065, Australia

⁵Department of Medicine, University of Melbourne, Parkville, VIC 2900, Australia

⁶Section of Transplantation, Department of Surgery, University of Chicago, 5841 South Maryland Avenue, Chicago, IL 60637, USA

⁷Mali International Center of Excellence in Research, University of Sciences, Techniques and Technologies of Bamako, 1805 Bamako, Mali

*Correspondence: mpsoares@igc.gulbenkian.pt

<http://dx.doi.org/10.1016/j.cell.2014.10.053>

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

SUMMARY

Glycosylation processes are under high natural selection pressure, presumably because these can modulate resistance to infection. Here, we asked whether inactivation of the UDP-galactose:β-galactoside-α1-3-galactosyltransferase ($\alpha 1,3GT$) gene, which ablated the expression of the Galα1-3Galβ1-4GlcNAc-R (α-gal) glycan and allowed for the production of anti-α-gal antibodies (Abs) in humans, confers protection against *Plasmodium* spp. infection, the causative agent of malaria and a major driving force in human evolution. We demonstrate that both *Plasmodium* spp. and the human gut pathobiont *E. coli* O86:B7 express α-gal and that anti-α-gal Abs are associated with protection against malaria transmission in humans as well as in α1,3GT-deficient mice, which produce protective anti-α-gal Abs when colonized by *E. coli* O86:B7. Anti-α-gal Abs target *Plasmodium* sporozoites for complement-mediated cytotoxicity in the skin, immediately after inoculation by *Anopheles* mosquitoes. Vaccination against α-gal confers sterile protection against malaria in mice, suggesting that a similar approach may reduce malaria transmission in humans.

INTRODUCTION

Humans have relatively high levels of circulating antibodies (Abs) recognizing xeno-glycans expressed by pathogens (Oyerlaren et al., 2009). As for other antigens, xeno-glycans cannot be targeted by the immune system when also expressed as

self-glycans. This limitation can be bypassed by natural selection of mutations that inactivate the expression of self-glycans (Bishop and Gagneux, 2007). Presumably, natural selection of such loss-of-function mutations tailored the human anti-glycan immune repertoire through evolution (Bishop and Gagneux, 2007). This notion is supported by the inactivation of the cytidine monophosphate-N-acetylneuraminic acid hydroxylase-like (*CMAH*) gene in humans, which suppressed the expression of N-glycolylneuraminic acid (Neu5Gc) (Hayakawa et al., 2001) and allowed for immune reactivity against Neu5Gc (Tangvoranuntakul et al., 2003). In a similar manner, inactivation of the α1,3GT gene, which suppressed the expression of the Galα1-3Galβ1-4GlcNAc-R (α-gal) carbohydrate in ancestral anthropoid primates that gave rise to humans (Galili and Swanson, 1991), also allowed for immune reactivity against α-gal (Galili et al., 1984). While it has been argued that this evolutionary process is driven to a large extent by the acquisition of immune-resistance against pathogens expressing such glycans (Bishop and Gagneux, 2007; Cywes-Bentley et al., 2013), this was never tested experimentally.

Humans do not express α-gal and up to 1%–5% of the repertoire of circulating immunoglobulin M (IgM) and immunoglobulin G (IgG) in healthy adults is directed against this glycan (Macher and Galili, 2008). Production of α-gal-specific Abs is thought to be driven by exposure to bacterial components of the microbiota expressing α-gal (Macher and Galili, 2008), including specific members of the *Klebsiella* spp., *Serratia* spp., and *Escherichia coli* spp. (Galili et al., 1988). Expression of α-gal by these *Enterobacteriaceae* is associated with the bacterial capsule and cell wall glycoproteins, as well as with lipopolysaccharide (LPS) (Galili et al., 1988). Gut colonization by the human pathobiont *E. coli* O86:B7 (Pal et al., 1969) recapitulates the etiology of anti-α-gal Ab production in mice (Posekany et al., 2002) and in primates (Mañez et al., 2001), as well as the production of Abs directed against the α-gal-related anti-B blood group glycan in chickens

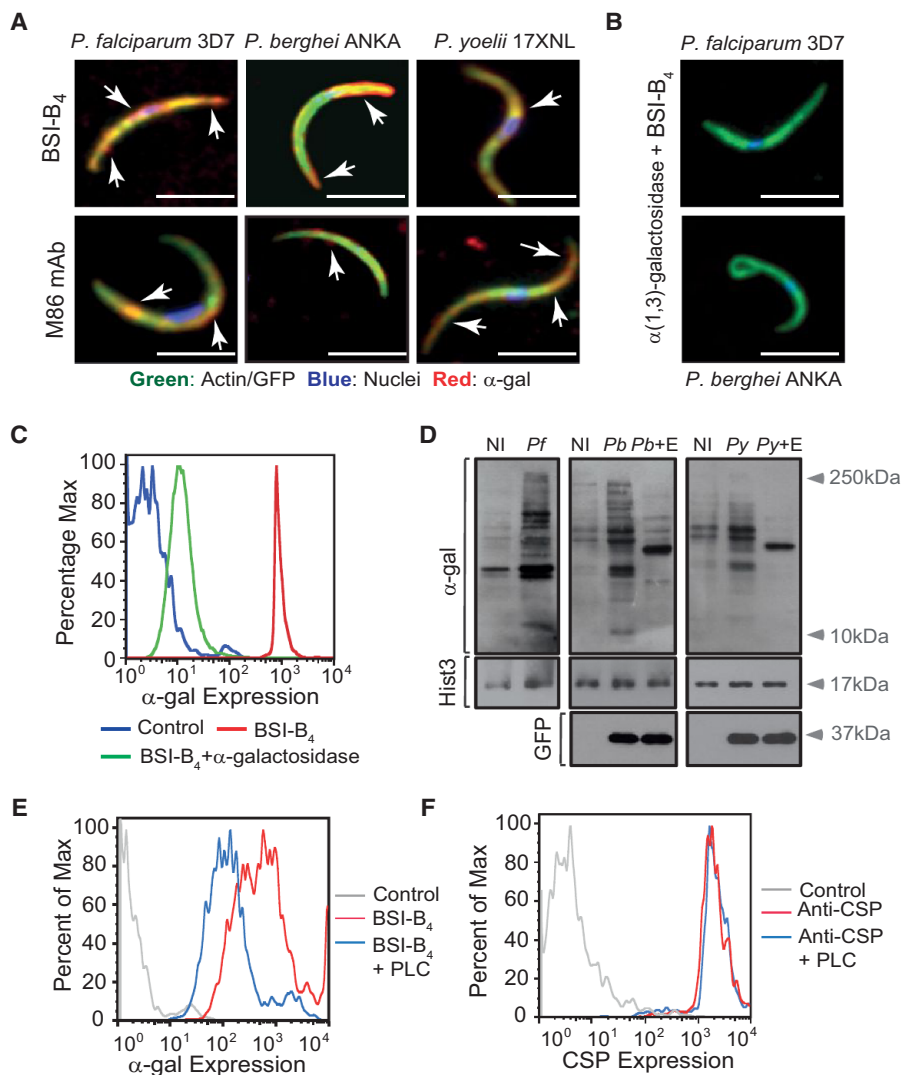


Figure 1. Detection of α -Gal in *Plasmodium* Sporozoites

(A) Composite images of GFP/actin (green), α -gal (red; white arrows), and DNA (blue) in *Plasmodium* sporozoites.

(B) Same staining as (A), after removal of α -gal by α -galactosidase. Images are representative of 2–3 independent experiments. Scale bar, 5 μ m.

(C) Detection of α -gal in *PbA*^{Hsp70-GFP} sporozoites by flow cytometry, representative of three independent experiments.

(D) Detection of α -gal in proteins extracted from salivary glands of noninfected (NI), *P. falciparum* 3D7 (Pf), *PbA*^{Hsp70-GFP} (Pb), or *P. yoelii* 17XNL (Py)-infected *A. mosquitoes*. Histone H3 (Hist3) and GFP were detected as loading controls. When indicated, α -gal was digested using α -galactosidase (E).

(E and F) Detection of α -gal (E) and CSP (F) in *PbA*^{Hsp70-GFP} sporozoites treated or not with phospholipase C (+PLC). Control is not stained. Data representative of 2–4 independent experiments.

See also Figure S1.

Malaria is transmitted to humans by the inoculation of *Plasmodium* sporozoites via the bite of female *Anopheles* (A.) mosquitoes (Ménard et al., 2013). While transmission may be rather efficient, only a fraction of the inoculated parasites manage to progress toward the establishment of infection (Rickman et al., 1990; Sauerwein et al., 2011; Verhage et al., 2005), hinting at a natural mechanism of protection that presumably targets the initial phases of the *Plasmodium* life cycle. Here, we demonstrate that production of anti- α -gal Abs in response to the gut *E. coli* O86:B7 pathobiont contributes

critically to this natural defense mechanism, reducing malaria transmission by *A. mosquitoes*.

RESULTS

Plasmodium spp. Express the α -Gal Glycan

The α -gal glycan was detected on the surface of *Plasmodium* sporozoites, as assessed by immunofluorescence for the human pathogen *Plasmodium falciparum* 3D7, as well as for the transgenic GFP-expressing strains of the rodent pathogens *Plasmodium berghei* ANKA (*PbA*) or *Plasmodium yoelii* 17XNL, using the lectin *Bandeiraea* (*Griffonia*) *simplicifolia*-I isolectin IB₄ (BSI-B₄) (Galili et al., 1985) or an anti- α -gal monoclonal antibody (M86 mAb) (Galili et al., 1998) (Figure 1A; Figures S1A and S1B available online). Specificity of α -gal detection was confirmed by its enzymatic removal using α -galactosidase (Figures 1B and 1C). Expression of α -gal was associated with proteins, as assessed by western blot in whole-cell extracts from *P. falciparum* 3D7, *PbA*, or *P. yoelii* 17XNL sporozoites

(Springer et al., 1959) and humans (Springer and Horton, 1969). This argues that gut colonization by *E. coli* O86:B7 may be particularly relevant in triggering the production of α -gal-specific Abs, presumably contributing to the high titers of these circulating Abs in healthy adult humans (Galili et al., 1988). Moreover, anti- α -gal Abs may also be produced in response to infection by pathogens expressing α -gal, such as illustrated for gram-negative bacteria from *Salmonella* spp. or for protozoan parasites from *Trypanosoma* spp. (Avila et al., 1989).

Anti- α -gal Abs are cytotoxic toward α -gal-expressing pathogens, as demonstrated in vitro for bacteria (Galili et al., 1988), protozoan parasites (Avila et al., 1989), and viruses enveloped by xenogeneic α -gal-expressing cell membranes (Takeuchi et al., 1996). Whether anti- α -gal Abs confer resistance to these and/or other pathogens in vivo has, to the best of our knowledge, not been established. Here, we tested this hypothesis specifically for *Plasmodium* spp. infection, the causative agent of malaria and a major driving force that shaped the evolution of anthropoid primates, including humans.

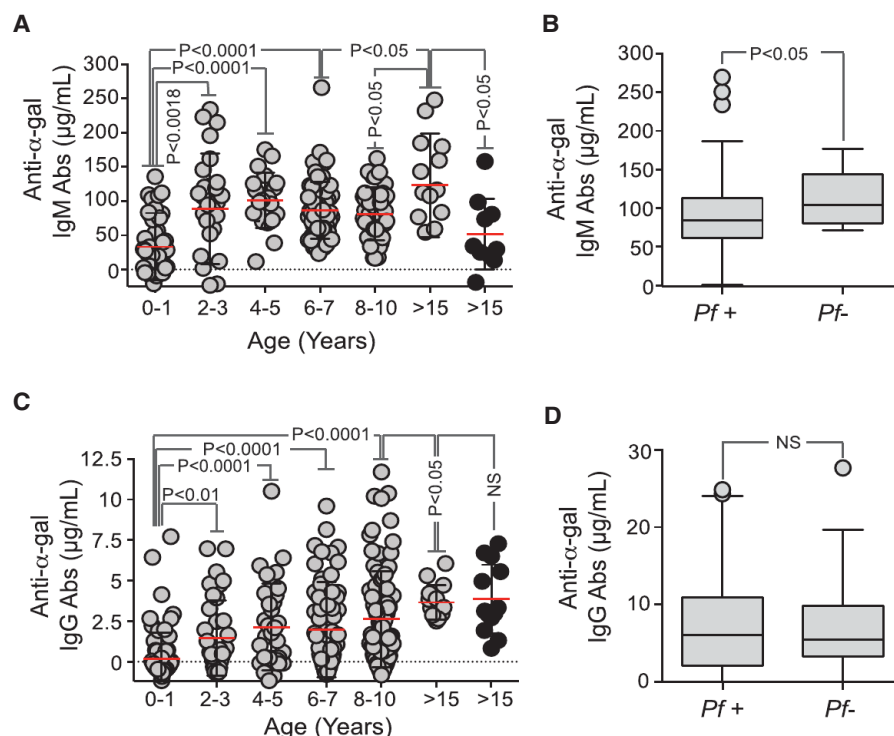


Figure 2. Anti- α -Gal IgM Abs Are Associated with Protection against Malaria Transmission in Individuals from a Malaria Endemic Region

(A) Anti- α -gal IgM Abs in individuals from a malaria endemic region in Mali (gray dots) or from the United States (black dots). Mean (red bars) \pm SD. (B) Levels of anti- α -gal IgM Abs in *P. falciparum*-infected (Pf+) versus noninfected (Pf-) children >4 years of age are shown as box plots in the same population as in (A).

(C) Anti- α -gal IgG Abs in individuals from a malaria endemic region in Mali (gray dots) or from the United States (black dots). Mean (red bars) \pm SD. (D) Levels of anti- α -gal IgG Abs in *P. falciparum*-infected (Pf+) versus noninfected (Pf-) children >4 years of age are shown as box plots in the same population as in (C).

(Figure 1D) and confirmed by enzymatic removal of α -gal (Figure 1D). Residual levels of α -gal were detected in the salivary glands of noninfected mosquitoes, suggesting that this glycan may be generated, at least partially, by *A. mosquitoes* (Figure 1D).

Expression of α -gal by *PbA* sporozoites was reduced by ~ 4 -fold when the glycosylphosphatidylinositol (GPI) anchor was cleaved by phospholipase C (PLC), as assessed by flow cytometry (Figure 1E). In contrast, GPI cleavage failed to reduce the expression of circumsporozoite protein (CSP), the main protein expressed at the surface of *Plasmodium* sporozoites (Figure 1F). This suggests that α -gal is bound to GPI-anchored surface proteins, including or not CSP, which despite being GPI-anchored (Moran and Caras, 1994) is resistant to PLC cleavage (Kimmel et al., 2003) (Figure 1F).

α -Gal-Specific IgM Abs Are Associated with Protection from *P. falciparum* Infection in Humans

We investigated whether a correlation exists between the levels of anti- α -gal Abs in healthy uninfected children and adults before the malaria season ($n = 330$ for IgG; $n = 229$ for IgM) and subsequent risk of *P. falciparum* infection (determined by biweekly PCR analysis of fingerprick blood samples) and febrile malaria (determined by weekly physical examination), during the ensuing 6 month malaria season in a cohort study in Mali, where this season is predictable and intense (Tran et al., 2014). In children <2 years, the average level of anti- α -gal IgM Abs was $33.4 \mu\text{g/mL}$ (95% confidence interval [CI]: 18.4 – $48.3 \mu\text{g/mL}$) (Figure 2A), similar to that reported in children with no history of malaria exposure (Avila et al., 1992; Doenz et al., 2000; Galili et al., 1984; Parker et al., 1999). However, anti- α -gal IgM Abs

anti- α -gal IgM Abs in children >4 years of age who had no *P. falciparum* infections detected during the 6-month malaria season ($n = 13$) was higher than those who became infected ($n = 141$) (Figure 2B). This suggests that there is a positive correlation between the levels of anti- α -gal IgM Abs and incidence of *P. falciparum* infection.

The average level of anti- α -gal IgG Abs in children <2 years was $1.46 \mu\text{g/mL}$ (95% CI: 0.22 – $0.69 \mu\text{g/mL}$) and increased in adults to $3.66 \mu\text{g/mL}$ (95% CI: 3.04 – $4.28 \mu\text{g/mL}$) (Figure 2C). In contrast to IgM, the levels of circulating α -gal-specific IgG were similar between malaria-exposed and nonexposed adults, suggesting that *P. falciparum* infection fails to drive an IgG response directed against α -gal (Figure 2D). This also suggests that there is no correlation between anti- α -gal IgG Abs and incidence of *P. falciparum* infection. Time-to-event analysis did not show a correlation between α -gal-specific IgM and IgG levels before the malaria season and subsequent risk of *P. falciparum* infection ($p = 0.76$ and $p = 0.08$, respectively) or febrile malaria ($p = 0.35$ and $p = 0.18$, respectively).

Gut Colonization by *E. coli* O86:B7 Elicits a Protective α -Gal-Specific IgM Ab Response against Malaria Transmission

To test whether anti- α -gal IgM Abs are protective against malaria transmission, we took advantage of “human-like” $\alpha 1,3\text{Gt}$ -deficient mice. Unlike humans, wild-type mice have a functional $\alpha 1,3\text{Gt}$ gene and express α -gal on secreted and cell-surface glycoconjugates, suppressing the development of anti- α -gal immunity (Yang et al., 1998). Deletion of $\alpha 1,3\text{Gt}$ gene eliminates α -gal (Tearle et al., 1996), allowing for anti- α -gal Ab production in $\alpha 1,3\text{Gt}^{-/-}$ mice (Chiang et al., 2000; Tearle et al., 1996; Yang

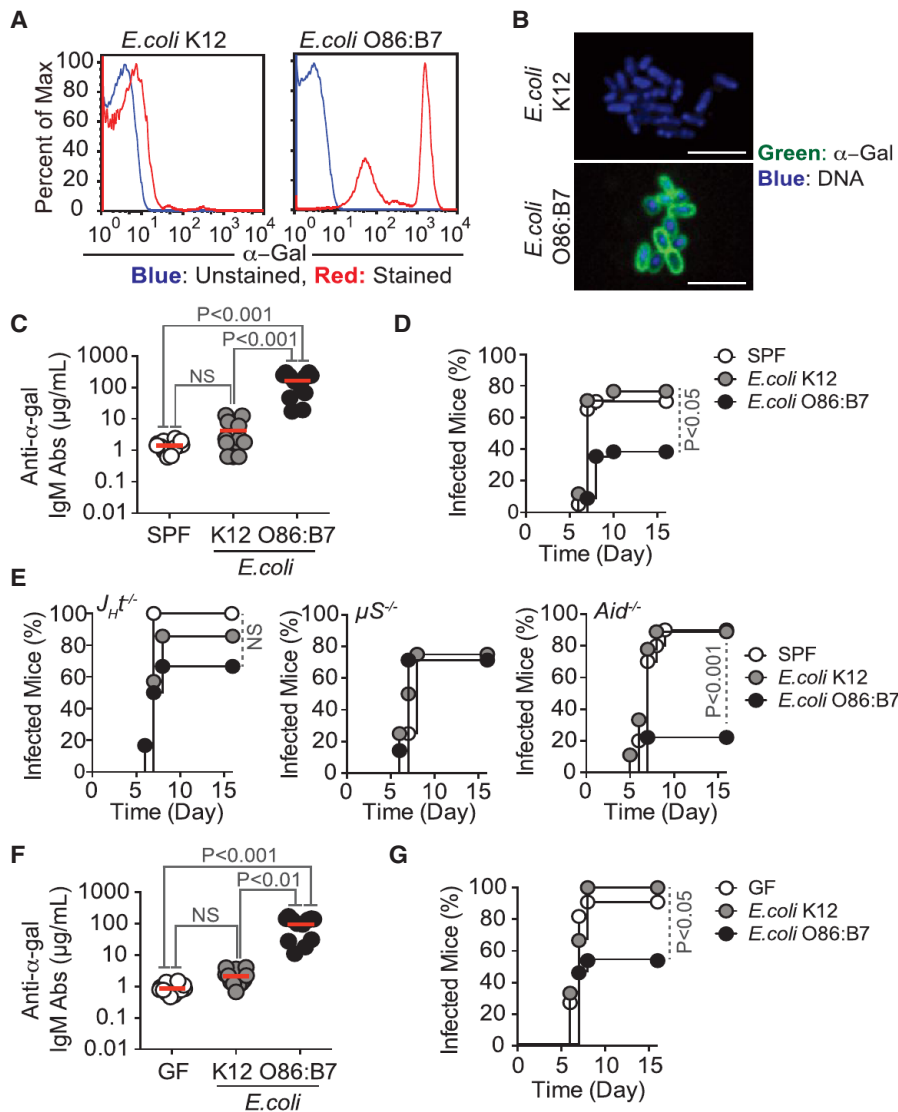


Figure 3. Gut Colonization by *E. coli* Expressing α -Gal Protects against *Plasmodium* Infection

(A and B) Detection of α -gal in *E. coli* strains by (A) flow cytometry and (B) immunofluorescence. Representative of 2–3 independent experiments. Composite images in (B), i.e., α -gal (green) and DNA (blue) at 100 \times magnification. Scale bar, 10 μ m.

(C and D) $\alpha 1,3Gt^{-/-}$ mice maintained under SPF were treated with streptomycin for 7 days. (C) Anti- α -gal IgM Abs levels were measured in $\alpha 1,3Gt^{-/-}$ mice not colonized (SPF), colonized with *E. coli* K12, or colonized with O86:B7 strains (2–3 experiments; $n = 12$). (D) Incidence of blood stage of *Plasmodium* infection (%) in mice colonized as in (C) and exposed to *PbA*^{EEF1a-GFP}-infected *A. stephensi* mosquitoes (four experiments; $n = 17$ –34).

(E) Incidence of blood stage of *Plasmodium* infection (%) in $\alpha 1,3Gt^{-/-}$ *J_HT^{-/-}*, $\alpha 1,3Gt^{-/-}$ *Aid^{-/-}*, and $\alpha 1,3Gt^{-/-}$ $\mu S^{-/-}$ mice colonized as in (C) and exposed to *PbA*^{Hsp70-GFP}-infected *A. stephensi* mosquitoes (1–2 experiments; $n = 4$ –10).

(F) Anti- α -gal IgM Abs were measured in GF $\alpha 1,3Gt^{-/-}$ mice not colonized (GF), colonized with *E. coli* K12, or colonized with O86:B7 strains (2–3 experiments; $n = 12$). (G) Incidence of blood stage of *Plasmodium* infection (%) in mice colonized as in (F) and exposed to *PbA*^{EEF1a-GFP}-infected *A. stephensi* mosquitoes (four experiments; $n = 9$ –13).

Mean (red bars).

See also Figure S2.

et al., 1998). However, $\alpha 1,3Gt^{-/-}$ mice are known to produce only residual levels of circulating anti- α -gal Abs when maintained under specific pathogen-free (SPF) conditions (Chiang et al., 2000). Production of anti- α -gal Abs can be enhanced upon enteric exposure to *E. coli* O86:B7 (Posekany et al., 2002). We confirmed that *E. coli* O86:B7 expresses high levels of α -gal (Yi et al., 2006), which is not the case for the *E. coli* K12 strain (Figures 3A and 3B). Colonization of $\alpha 1,3Gt^{-/-}$ mice by *E. coli* O86:B7 after antibiotic treatment (streptomycin sulfate; 5 g/l in drinking water for 7 days prior to colonization) increased the levels of circulating anti- α -gal IgM Abs from 1.4 μ g/ml (95% CI: 1.1–1.8 μ g/ml) to 162.9 μ g/ml (95% CI: 95.89–230.1 μ g/ml) before and after colonization, respectively (Figure 3C). Levels of anti- α -gal IgM Abs in colonized $\alpha 1,3Gt^{-/-}$ mice were in the range of adult individuals from a malaria endemic region (Figure 2A). In contrast, the levels of circulating anti- α -gal IgG Abs remained at residual levels, i.e., <1 μ g/ml (Figure S2A), again in the range of adult individuals from a malaria endemic region (Fig-

ure 2C). Colonization by *E. coli* K12 did not induce the production of circulating anti- α -gal Abs (Figure 3C). Gut colonization by *E. coli* O86:B7 was associated with protection of $\alpha 1,3Gt^{-/-}$ mice from *PbA* transmission by infected *Anopheles stephensi* mosquitoes (Figure 3D). This

was not the case when $\alpha 1,3Gt^{-/-}$ mice were or were not colonized by *E. coli* K12 (Figure 3D). To determine whether the protective effect associated with gut colonization by *E. coli* O86:B7 is mediated by anti- α -gal Abs, we performed similar colonization experiments in $\alpha 1,3Gt^{-/-}$ *J_HT^{-/-}* lacking B cells (Gu et al., 1993), $\alpha 1,3Gt^{-/-}$ $\mu S^{-/-}$ mice lacking circulating IgM (Ehrenstein et al., 1998) or $\alpha 1,3Gt^{-/-}$ *Aid^{-/-}* mice that fail to undergo Ig class switch recombination or somatic hypermutation (Muramatsu et al., 2000). Gut colonization by *E. coli* O86:B7 failed to protect $\alpha 1,3Gt^{-/-}$ *J_HT^{-/-}* and $\alpha 1,3Gt^{-/-}$ $\mu S^{-/-}$, but not $\alpha 1,3Gt^{-/-}$ *Aid^{-/-}* mice from *PbA*-infected mosquitoes, as compared to genetic-matched control mice colonized or not by *E. coli* K12 (Figure 3E). This shows that the protective effect of gut colonization by *E. coli* O86:B7 acts via a mechanism mediated by anti- α -gal IgM Abs that do not undergo somatic hypermutation.

Germ-free (GF) $\alpha 1,3Gt^{-/-}$ mice had low but detectable levels of anti- α -gal IgM Abs, i.e., 0.87 μ g/ml (95% CI: 0.66–1.1 μ g/ml),

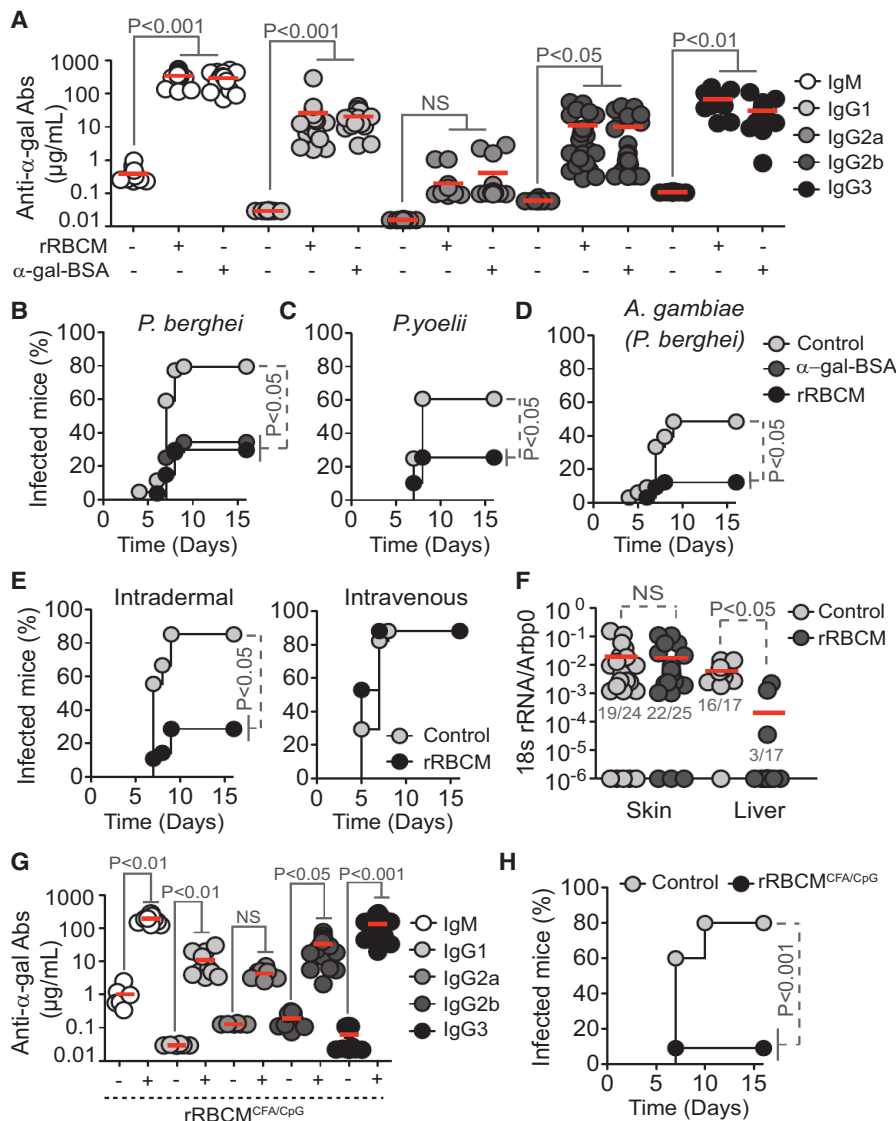


Figure 4. Protective Effect of α -Gal Immunization

(A) Anti- α -gal Abs in the serum of control (–) versus rRBCM (+) or α -gal-BSA (+) immunized $\alpha 1,3Gt^{-/-}$ mice (2–3 experiments; $n = 12$ –29).

(B–D) Incidence of blood stage of infection (%) in $\alpha 1,3Gt^{-/-}$ mice treated as in (A) and exposed to (B) *PbA*^{EEF1a-GFP}-infected *A. stephensi* mosquitoes (seven experiments; $n = 27$ –44), (C) *P. yoelii* 17XNL-infected *A. stephensi* mosquitoes (five experiments; $n = 28$ –39), or (D) *PbA*^{EEF1a-GFP}-infected *A. gambiae* mosquitoes (four experiments; $n = 27$ –34).

(E) Incidence of blood stage of infection (%) in nonimmunized (control) versus immunized (rRBCM) $\alpha 1,3Gt^{-/-}$ mice receiving *PbA*^{EEF1a-GFP} sporozoites (3–4 experiments; $n = 17$ –28).

(F) *Plasmodium* 18 s rRNA/Arbp0 mRNA in skin and liver of nonimmunized (control) versus immunized (rRBCM) $\alpha 1,3Gt^{-/-}$ mice exposed to *PbA*^{EEF1a-GFP}-infected *A. stephensi* mosquitoes (3–5 experiments). Infected/total mice (gray nbs).

(G) Same as (A) in control (–) versus immunized (+; rRBCM emulsified in CFA+CpG) $\alpha 1,3Gt^{-/-}$ mice (two experiments; $n = 6$ –23).

(H) Incidence of blood stage of infection (%) in $\alpha 1,3Gt^{-/-}$ mice treated as in (G) and infected as in (B) (three experiments; $n = 16$ –19). In (A), (F), and (G), dots are individual mice and mean (red bars). See also Figures S3 and S4.

suggesting that these are natural Abs (Figure 3F). The production of anti- α -gal IgM Abs in GF $\alpha 1,3Gt^{-/-}$ mice being driven by expression of these glycans in food components is possible, but this has not been tested. GF $\alpha 1,3Gt^{-/-}$ mice did not produce anti- α -gal IgG Abs (Figure S2C). Susceptibility to *PbA* transmission by infected *A. mosquitoes* was similar in SPF versus GF $\alpha 1,3Gt^{-/-}$ mice (Figures 3D and 3G). When GF $\alpha 1,3Gt^{-/-}$ mice were monocolonized by *E. coli* O86:B7, the levels of circulating anti- α -gal IgM Abs increased to 96.62 μ g/ml (95% CI: 59.32–133.9 μ g/ml) (Figure 3F), which is the range in which adult individuals from a malaria endemic region (Figure 2A), without concomitant induction of anti- α -gal IgG Abs (Figure S2C). Monocolonization by *E. coli* O86:B7, but not by *E. coli* K12, protected $\alpha 1,3Gt^{-/-}$ mice from *PbA* transmission by *A. mosquitoes* (Figure 3F). This suggests that gut colonization by a specific pathobiont expressing α -gal recapitulates to a large extent the normal etiology of the human anti- α -gal Ab response (Figure 2) and in-

duces protection against *Plasmodium* infection, such as observed in a malaria endemic region (Figure 2).

It should be noted that the percentage of infected red blood cell (RBC), i.e., parasitemia, and incidence of mortality were similar among those $\alpha 1,3Gt^{-/-}$ mice that were infected by *PbA* regardless of colonization (Figures S2B and S2D). This suggests that gut colonization by *E. coli* O86:B7 protects against *Plasmodium* transmission, but not against disease once the erythrocytic stage of infection is established.

Immunization against α -Gal Protects from *Plasmodium* Transmission

Immunization of $\alpha 1,3Gt^{-/-}$ mice against α -gal, using rabbit RBC membranes (rRBCM) expressing high levels of α -gal or synthetic α -gal conjugated to BSA (α -gal-BSA) elicited the production of circulating anti- α -gal IgM and IgG Abs (Figure 4A). Control $\alpha 1,3Gt^{+/+}$ mice failed to produce anti- α -gal Abs (Chiang et al., 2000; Tearle et al., 1996; Yang et al., 1998) (Figure S3A). Circulating anti- α -gal immunoglobulin A (IgA) and immunoglobulin E (IgE) Abs were undetectable in control or immunized $\alpha 1,3Gt^{-/-}$ and $\alpha 1,3Gt^{+/+}$ mice (data not shown). The concentration of anti- α -gal IgM Abs in the plasma of immunized $\alpha 1,3Gt^{-/-}$ mice was in the range of adult individuals from malaria endemic regions (Figure 2A). Circulating anti- α -gal IgG Abs in immunized

$\alpha 1,3Gt^{-/-}$ mice, predominantly from IgG1, IgG2b, and IgG3 subclasses, were present at higher concentrations, as compared to total IgG in adult individuals from malaria endemic regions. Little or no circulating IgG2a (Figure 4A) or IgG2c (data not shown) were detected in immunized $\alpha 1,3Gt^{-/-}$ mice.

Immunization against α -gal protected $\alpha 1,3Gt^{-/-}$ mice from *PbA* (Figure 4B) and *P. yoelli* 17XNL (Figure 4C) transmission by infected *A. stephensi* mosquitoes, as well as from *PbA* transmission by *A. gambiae* mosquitoes (Figure 4D) versus control nonimmunized $\alpha 1,3Gt^{-/-}$ mice. Control immunized $\alpha 1,3Gt^{+/+}$ mice were neither protected from *PbA* (Figure S3B) nor *P. yoelli* 17XNL (Figure S3C) transmission by *A. stephensi* mosquitoes nor against *PbA* transmission by *A. gambiae* mosquitoes (Figure S3D) versus naive $\alpha 1,3Gt^{+/+}$ mice.

Immunized $\alpha 1,3Gt^{-/-}$ mice were protected from artificial transmission of *PbA* sporozoites via intradermal inoculation versus control nonimmunized $\alpha 1,3Gt^{-/-}$ mice (Figure 4E) or control immunized or nonimmunized $\alpha 1,3Gt^{+/+}$ mice (Figure S3E). Protection was no longer observed when sporozoites were inoculated intravenously (Figures 4E and S3E). This suggests that the protective effect of α -gal immunization is exerted in the dermis, presumably via an anti- α -gal Ab driven mechanism that is no longer effective once sporozoites reach the blood.

PbA transmission was associated with accumulation of *Plasmodium* 18S rRNA at the site of inoculation, as quantified in the ear pinna by qRT-PCR (Figure 4F). The relative amount of *Plasmodium* 18S rRNA was similar in immunized versus control nonimmunized $\alpha 1,3Gt^{-/-}$ mice (Figure 4F) or control immunized or nonimmunized $\alpha 1,3Gt^{+/+}$ mice (Figure S3F). Immunized $\alpha 1,3Gt^{-/-}$ mice did not accumulate *Plasmodium* 18S rRNA in the liver, when compared to control nonimmunized $\alpha 1,3Gt^{-/-}$ mice (Figure 4F) or control nonimmunized or immunized $\alpha 1,3Gt^{+/+}$ mice (Figure S3F). This suggests that α -gal immunization arrests the transit of inoculated sporozoites from the skin into the liver, without interfering with sporozoite inoculation by *A. mosquitoes*.

TLR9 Agonist Adjuvant Enhances the Protective Effect of α -Gal Immunization

Immunization of $\alpha 1,3Gt^{-/-}$ mice with rRBCM emulsified in complete Freund's adjuvant (CFA), supplemented with toll-like receptor 9 agonist CpG, enhanced anti- α -gal IgG2b and IgG3 Ab response by 2- to 3-fold (Figure 4G) versus immunization without adjuvant (Figure 4A). This was associated with 88% reduction in the relative risk of transmission of *PbA* infection by *A. mosquitoes* (95% CI: 0.032–0.452) versus 61% reduction upon immunization without adjuvant (95% CI: 0.209–0.726) (Figures 4B and 4H). This protective effect was not observed in control $\alpha 1,3Gt^{+/+}$ mice (Figures S3G and S3H).

Parasitemias were similar in immunized $\alpha 1,3Gt^{-/-}$ mice not protected from *PbA* infection versus control nonimmunized $\alpha 1,3Gt^{-/-}$ mice as well as control nonimmunized or immunized $\alpha 1,3Gt^{+/+}$ mice (data not shown). Moreover, when infected, all mice succumbed to experimental cerebral malaria. This suggests that while protective against malaria transmission, α -gal immunization is not protective against the development of severe disease if *Plasmodium* manages to establish infection. In

keeping with this notion, when inoculated with *PbA*-infected RBC, immunized $\alpha 1,3Gt^{-/-}$ mice developed similar levels of parasitemia and disease severity, as compared to control nonimmunized $\alpha 1,3Gt^{-/-}$ mice as well as to control nonimmunized or immunized $\alpha 1,3Gt^{+/+}$ mice (Figure S4A).

We tested further whether the protective effect conferred by α -gal immunization is associated with sterile protection, i.e., inability of *Plasmodium* to establish blood stage of infection. Passive transfer of RBCs harvested from protected immunized $\alpha 1,3Gt^{-/-}$ mice at day 8–9 post-*PbA* transmission by *A. mosquitoes* failed to transmit disease to naive $\alpha 1,3Gt^{-/-}$ mice (Figure S4B). In contrast, passive transfer of RBC harvested from nonprotected immunized $\alpha 1,3Gt^{-/-}$ mice, readily transmitted disease to naive $\alpha 1,3Gt^{-/-}$ mice (Figure S4B). This demonstrates that the protective effect of immunization against α -gal is associated with sterile protection against malaria.

Anti- α -Gal IgM and IgG Abs Produced in Response to α -Gal Immunization Confer Protection against Malaria Transmission

We asked whether the protective effect of α -gal immunization is mediated by anti- α -gal IgM and/or IgG Abs. Immunized $\alpha 1,3Gt^{-/-}J_HT^{-/-}$ mice failed to produce anti- α -gal IgM or IgG Abs versus naive $\alpha 1,3Gt^{-/-}J_HT^{-/-}$ mice or immunized $\alpha 1,3Gt^{-/-}$ mice (Figure 5A). Moreover, immunized $\alpha 1,3Gt^{-/-}J_HT^{-/-}$ mice were not protected against *PbA* transmission by *A. mosquitoes* versus control nonimmunized $\alpha 1,3Gt^{-/-}J_HT^{-/-}$ mice (Figure 5B). This shows that the protective effect of α -gal immunization is mediated via a B cell-dependent mechanism.

Immunization of $\alpha 1,3Gt^{-/-}Aid^{-/-}$ mice failed to induce the production of anti- α -gal IgG, but not IgM Abs, versus naive $\alpha 1,3Gt^{-/-}Aid^{-/-}$ or immunized $\alpha 1,3Gt^{-/-}$ mice (Figure 5A). Immunized $\alpha 1,3Gt^{-/-}Aid^{-/-}$ mice were nevertheless protected against *PbA* transmission by *A. mosquitoes* versus nonimmunized $\alpha 1,3Gt^{-/-}Aid^{-/-}$ mice (Figure 5B). This confirms that anti- α -gal IgM Abs can confer protection against malaria transmission (Figure 2B) and that the protective effect of α -gal-specific IgM Abs does not require somatic hypermutation.

Immunization of $\alpha 1,3Gt^{-/-}\mu S^{-/-}$ mice failed to induce anti- α -gal IgM Abs, without interfering with anti- α -gal IgG Ab response versus naive $\alpha 1,3Gt^{-/-}\mu S^{-/-}$ mice or immunized $\alpha 1,3Gt^{-/-}$ mice (Figure 5A). Immunized $\alpha 1,3Gt^{-/-}\mu S^{-/-}$ mice were nevertheless protected from *PbA* transmission by *A. mosquitoes* versus control naive $\alpha 1,3Gt^{-/-}\mu S^{-/-}$ mice (Figure 5B). Immunized $\alpha 1,3Gt^{-/-}$ mice did not produce circulating anti- α -gal IgA or IgE Abs (data not shown) and a putative protective effect for these Ig isotypes was excluded. This demonstrates that anti- α -gal IgG Abs produced in response to immunization confer protection against malaria transmission.

Immunization of $\alpha 1,3Gt^{-/-}Tcr\beta^{-/-}$ mice lacking mature $\alpha\beta$ T cells (Mombaerts et al., 1992) compromised anti- α -gal IgM and IgG response versus control immunized $\alpha 1,3Gt^{-/-}$ mice (Figure 5A). Immunized $\alpha 1,3Gt^{-/-}Tcr\beta^{-/-}$ mice were not protected from *PbA* transmission by *A. mosquitoes* versus control naive $\alpha 1,3Gt^{-/-}Tcr\beta^{-/-}$ mice (Figure 5B). This shows that anti- α -gal Abs produced in response to immunization are T cell-dependent (Cretin et al., 2002) and so is their protective effect.

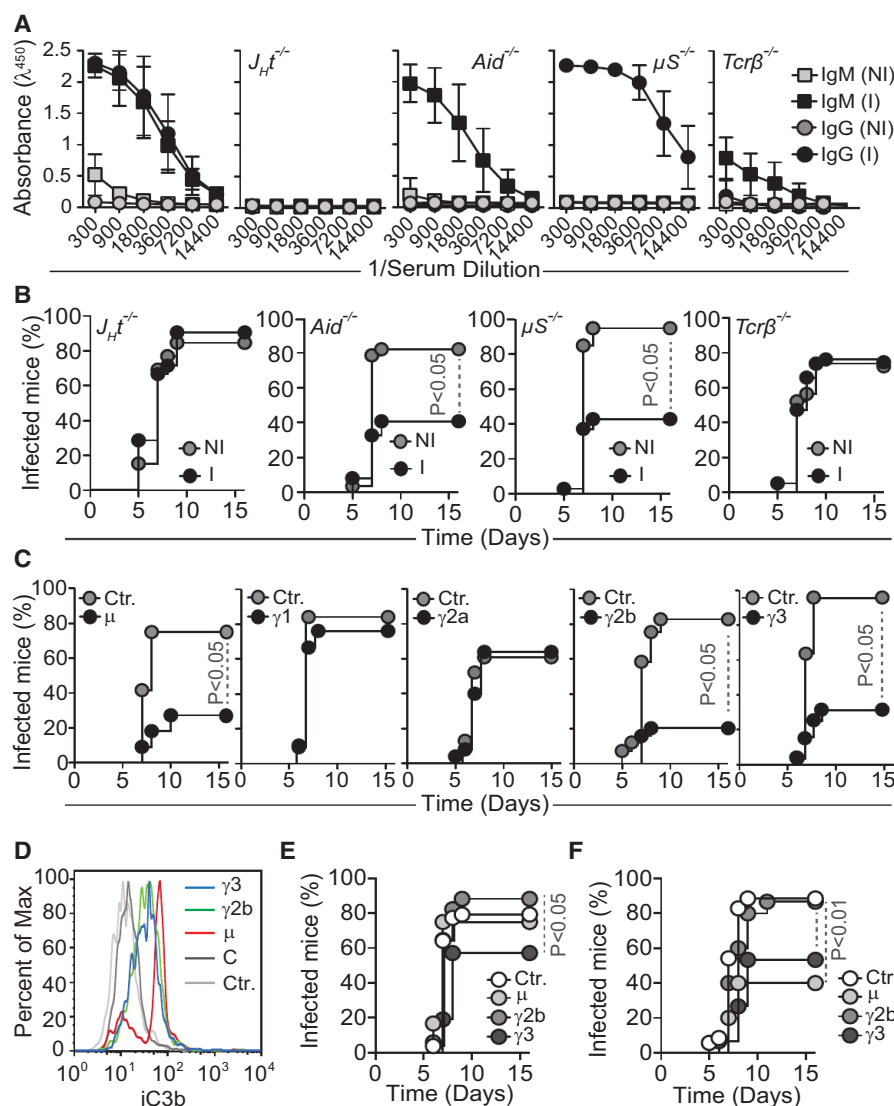


Figure 5. Protective Effect of Anti- α -gal Abs

(A) Relative absorbance of anti- α -gal Abs (Mean \pm SD) in serial serum dilutions from nonimmunized (NI) or rRBCM-immunized (I) $\alpha 1,3Gt^{-/-}$ mice (two experiments; $n = 10$).

(B) Incidence of blood stage infection (%) in specific immune component-deleted $\alpha 1,3Gt^{-/-}$ mice immunized (I) or not (NI) as in (A) and exposed to *PbA*^{EEF1a-GFP}-infected mosquitoes (3–7 experiments; $n = 13$ –41).

(C) Incidence of blood stage of infection (%) in $\alpha 1,3Gt^{-/-}$ mice after passive transfer of anti- α -gal Abs versus controls (no passive transfer; ctr.) exposed to *PbA*^{EEF1a-GFP}-infected mosquitoes (4–7 experiments; $n = 19$ –32).

(D) C3 deposition in *PbA*^{Hsp70-GFP} sporozoites not exposed (ctr.) or exposed to anti- α -gal Abs plus mouse complement (C). Representative of three independent experiments.

(E) Incidence of blood stage of infection (%) in $\alpha 1,3Gt^{-/-}$ C3^{-/-} mice after passive transfer of anti- α -gal IgM (μ), IgG2b ($\gamma 2b$), or IgG3 ($\gamma 3$) Abs versus controls (ctr.; no passive transfer) not receiving Abs, exposed to *PbA*^{EEF1a-GFP}-infected mosquitoes (four experiments; $n = 21$ –37).

(F) Same as (E) in PMN-depleted $\alpha 1,3Gt^{-/-}$ mice (four experiments; $n = 15$ –25).

See also Figures S5 and S6.

Naive and immunized $\alpha 1,3Gt^{-/-}$ $J_H T^{-/-}$ (Figure S5A), $\alpha 1,3Gt^{-/-}$ $Aid^{-/-}$ (Figure S5B) and $\alpha 1,3Gt^{-/-}$ $\mu S^{-/-}$ (Figure S5C) mice, not protected from *PbA* transmission, developed similar levels of parasitemia and succumbed to experimental cerebral malaria. This was not the case for $\alpha 1,3Gt^{-/-}$ $Tcr\beta^{-/-}$ mice (Figure S5D), consistent with the involvement of T cells in the pathogenesis of experimental cerebral malaria (Belnoue et al., 2002).

Passive transfer of anti- α -gal IgM to naive $\alpha 1,3Gt^{-/-}$ mice conferred protection against *PbA* transmission by *A. mosquitoes* (Figure 5C). This was also the case for passive transfer of anti- α -gal Abs from specific IgG subclasses, namely, IgG2b and IgG3 (Figure 5C), but not IgG1 or IgG2a (Figure 5C). Relative binding to α -gal was similar for all mAbs tested, as assessed by ELISA using α -gal-BSA as a solid-phase antigen (Figure S6A) or by immunofluorescence using *PbA* sporozoites (Figure S6B). Specificity of anti- α -gal binding to *Plasmodium* sporozoites was assessed by enzymatic removal of α -gal, confirming that these mAbs recognize specifically and only the α -gal glycan on the sur-

face of *Plasmodium* sporozoites (Figure S6C). IgG2a, IgG2b, and IgG3 mAbs are class-switched mutants derived from the original anti- α -gal IgG1 clone and as such have similar affinities for α -gal (Ding et al., 2008). These data reveal that while IgM anti- α -gal Abs are sufficient per se to confer protection against malaria transmission, this protective effect can be enhanced when specific subclasses anti- α -gal IgG Abs are present at sufficient high levels.

Once bound to the surface of *Plasmodium* sporozoites, anti- α -gal IgM, IgG2b, and IgG3 mAbs activated the classical pathway of complement, as assessed by C3 deposition (Figure 5D). Anti- α -gal IgG1 or IgG2a mAbs failed to activate complement (data not shown), and complement activation was also not observed in the absence of anti- α -gal Abs (Figures 5D and S7), showing that the alternative and lectin pathways of complement are not activated by *Plasmodium* sporozoites.

We then asked whether the protective effect exerted by anti- α -gal Abs is mediated via a mechanism involving the activation of the complement cascade (Figure 5D) (Ding et al., 2008; Miyatake et al., 1998). Passive transfer of anti- α -gal IgM Abs or anti- α -gal IgG2b mAb to $\alpha 1,3Gt^{-/-}$ C3^{-/-} mice, which lack C3 and cannot activate the complement cascade, failed to confer protection against *PbA* transmission versus control $\alpha 1,3Gt^{-/-}$ C3^{-/-} mice (Figure 5E). Passive transfer of anti- α -gal IgG3 mAb to $\alpha 1,3Gt^{-/-}$ C3^{-/-} mice conferred residual but significant protection versus control $\alpha 1,3Gt^{-/-}$ C3^{-/-} mice (Figure 5E). This

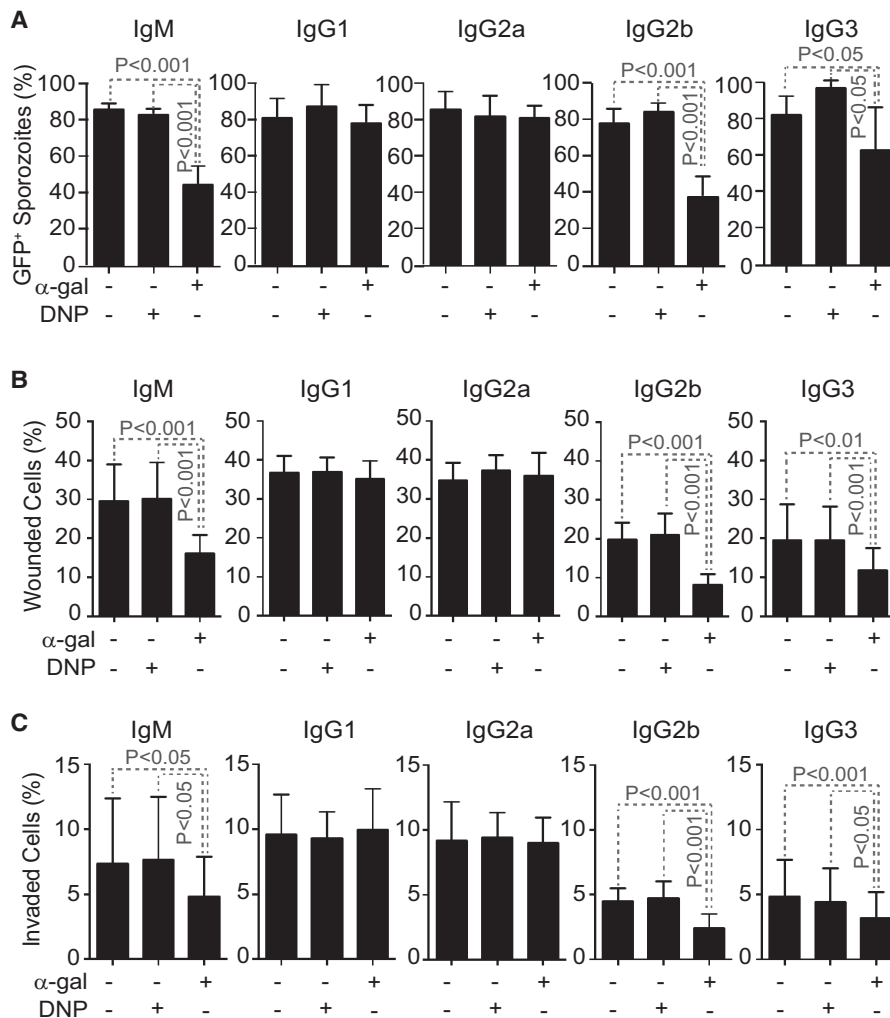


Figure 6. Protective Effect of Anti- α -Gal Abs against Hepatocyte Infection

(A) Mean percentage (%) of viable GFP⁺ *PbA*^{Hsp70-GFP} sporozoites \pm STD (3–4 experiments) after exposure in vitro to anti- α -gal or control anti-DNP mAbs in the presence of mouse complement. (B and C) Mean percentage (%) of HepG2 cells (B) wounded (Dextran-Red⁺) or (C) invaded (GFP⁺) by *PbA*^{Hsp70-GFP} sporozoites treated as in (A) \pm SD (six experiments).

while preventing the onset of cerebral malaria (data not shown), consistent with previous findings (Chen et al., 2000).

Anti- α -Gal Abs Are Cytotoxic to *Plasmodium* Sporozoites

Complement activation by anti- α -gal IgM, IgG2b, or IgG3 mAb was cytotoxic to *PbA* sporozoites in vitro, as assessed by sporozoite GFP expression (Figure 6A). Anti- α -gal IgG1 and IgG2a mAbs, which did not activate complement when bound to *Plasmodium* sporozoites (data not shown), were not cytotoxic (Figure 6A). The cytotoxic effect of anti- α -gal IgM, IgG2b, and IgG3 was similar when using mouse (Figure 6A) or rabbit (Figure S7A) complement but was strictly dependent on the presence of complement (Figure S7B). A similar cytotoxic effect was observed when quantifying viable “crescent-shaped” sporozoites (Figure S7C), an independent readout for sporozoite viability (Hegge et al., 2010). Isotype-matched control anti-dinitrophenyl (DNP)

Abs were not cytotoxic to *PbA* sporozoites in vitro (Figures 6A and S7A–S7C).

Anti- α -Gal Abs Inhibit Hepatocyte Invasion by *Plasmodium* Sporozoites

We asked whether anti- α -gal Abs inhibit hepatocyte transmigration (wounding) and/or hepatocyte invasion by *Plasmodium* sporozoites (Mota et al., 2001). Complement activation by anti- α -gal IgM, IgG2b, and IgG3 Abs inhibited hepatocyte transmigration (Figure 6B) and invasion (Figure 6C), as assessed in vitro for *PbA* sporozoites. This inhibitory effect was not observed when using anti- α -gal IgG1 or IgG2a Abs or isotype/subclass-matched control anti-DNP Abs (Figures 6B and 6C).

We then assessed whether anti- α -gal Abs inhibit the development of exoerythrocytic forms (EEF) of *Plasmodium*. Complement activation by anti- α -gal IgM, IgG2b, and IgG3 Abs reduced the number of EEF (Figure 7A), as well as the average EEF size (Figures 7B and 7C) formed in vitro by *PbA* sporozoites. Anti- α -gal IgG1 Abs did not show this inhibitory effect, while anti- α -gal IgG2a Abs did not reduce the number of EEF (Figure 7A) but had a residual inhibitory effect on EEF size (Figures 7B and

suggests that the protective effect exerted by anti- α -gal IgM and IgG2b Abs acts via a mechanism that is strictly complement dependent, whereas the protective effect of anti- α -gal IgG3 Abs is partially but probably not strictly dependent on complement activation. Infection incidence was similar in control $\alpha 1,3Gt^{-/-}$ C3^{-/-} versus $\alpha 1,3Gt^{-/-}$ C3^{+/+} mice (Figures 5C and 5E).

Complement activation generates C3a and C5a chemoattractants that promote IgG-dependent polymorphonuclear (PMN) cell cytotoxicity (Ding et al., 2008; Nimmerjahn and Ravetch, 2008; Yin et al., 2004). Therefore, we asked whether the protective effect of anti- α -gal Abs involves PMN cells. Passive transfer of anti- α -gal IgG2b Abs to $\alpha 1,3Gt^{-/-}$ mice, depleted from PMN cells by the administration of anti-Ly-6G (Gr-1) (Porcherie et al., 2011), failed to confer protection against *PbA* transmission by *A. mosquitoes*, whereas passive transfer of anti- α -gal IgM or IgG3 Abs conferred protection (Figure 5F). This suggests that the protective effect exerted by anti- α -gal IgM and IgG2b Abs acts via a mechanism strictly dependent on PMN cells, whereas the protective effect of anti- α -gal IgG3 Abs is partially but probably not strictly dependent on PMN cells. Depletion of PMN cells per se did not interfere with *Plasmodium* infection (Figure 5F)

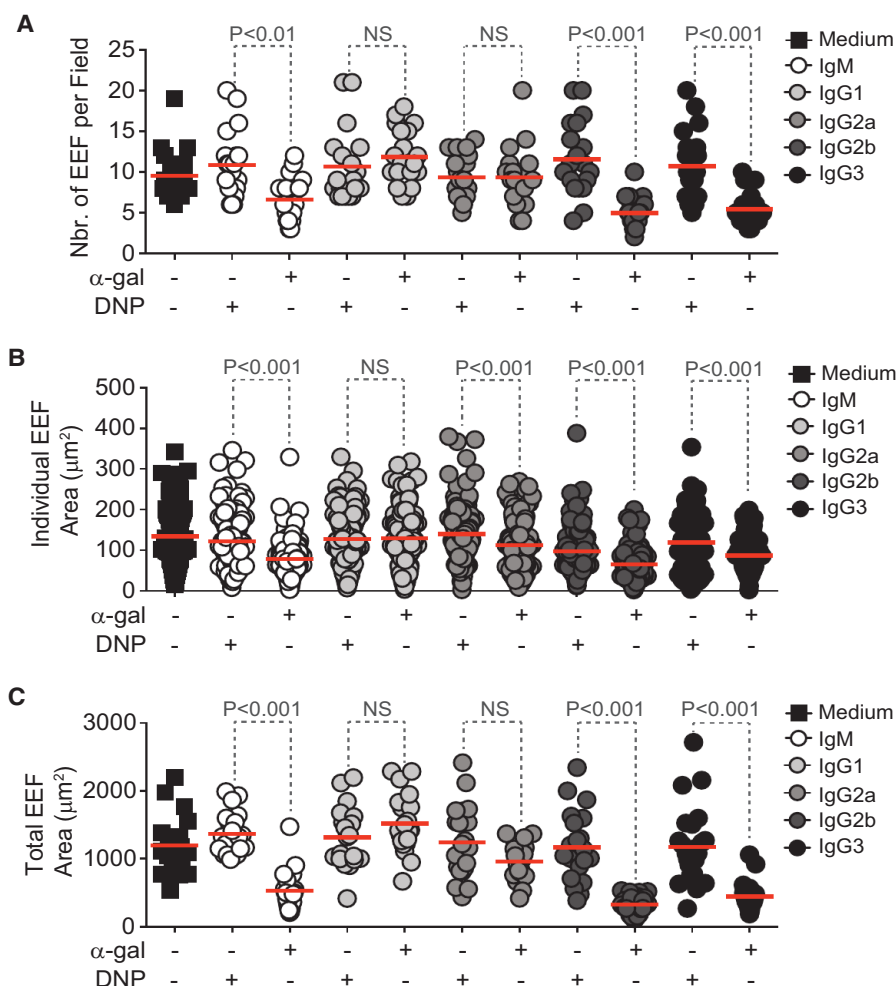


Figure 7. Protective Effect of Anti- α -Gal Abs against *Plasmodium* Maturation in Hepatocytes

(A) Number of EEF per field (dots; 20–23 fields). (B) Area of individual EEF (dots) ($n = 111$ –256 EEFs counted in 20–23 fields). (C) Total area of EEF (dots) per field (20–23 fields). HepG2 cells were incubated with *PbA*^{Hsp70-GFP} sporozoites, previously exposed to anti- α -gal or control anti-DNP mAbs in the presence of complement (A–C). See also Figure S7.

The protective effect exerted by anti- α -gal IgM Abs should be relevant to understand why malaria incidence is higher in children versus adults from malaria endemic regions (Modiano et al., 1996). Relative absence of these antibodies in children under the age of 2–3 years should favor malaria transmission, as compared to adults that have higher levels of circulating anti- α -gal IgM Abs (Figure 2A). This relative absence of anti- α -gal IgM in children may be explained by the (1) kinetics of establishment of an adult-like gut microbiota (Ringel-Kulka et al., 2013), (2) requirement of environmental and dietary exposure in the establishment of an adult-like gut microbiota, and/or (3) the kinetics of the establishment of adult-like B cell repertoire, including anti- α -gal B cells.

The protective effect exerted by anti- α -gal IgM Abs might also contribute to

7C). Isotype/subclass-matched control anti-DNP Abs did not modulate EEF numbers (Figure 7A) or average size (Figures 7B and 7C).

DISCUSSION

When inoculated in humans through the bite of an *A. mosquito*, *Plasmodium* sporozoites are confronted with relatively high levels of cytotoxic anti- α -gal IgM Abs (Figure 2A). That these Abs are protective against malaria transmission is supported by three independent lines of evidence. First, individuals from a malaria endemic region that show evidence of decreased *P. falciparum* infection risk have higher levels of circulating α -gal-specific IgM Abs, as compared to individuals who are susceptible to *P. falciparum* infection (Figure 2B). Second, when present at levels similar to those observed in individuals from a malaria endemic region—in $\alpha 1,3\text{Gt}^{-/-}$ mice colonized by human gut pathobiont *E. coli* O86:B7 expressing α -gal (Figure 3) or in immunized $\alpha 1,3\text{Gt}^{-/-}$ mice (Figures 4A and 4B)—anti- α -gal IgM Abs confer protection against malaria transmission. Third, passive transfer of anti- α -gal IgM Abs is sufficient per se to protect $\alpha 1,3\text{Gt}^{-/-}$ mice from malaria transmission (Figure 5C).

explain why only a small fraction of *Plasmodium* sporozoites inoculated by mosquitoes manage to progress toward the establishment of infection in humans. This is true even when *Plasmodium* sporozoites are inoculated under controlled experimental conditions in adults (Rickman et al., 1990; Sauerwein et al., 2011; Verhage et al., 2005). Presumably, when present at sufficient high levels in adults, circulating anti- α -gal IgM Abs prevent the large majority of *Plasmodium* sporozoites from establishing a successful infection. However, infection is established if as few as a couple of *Plasmodium* sporozoites manage to escape this natural mechanism of protection.

Whether α -gal detected at the surface of *Plasmodium* sporozoites (Figure 1) is produced by *Plasmodium* and/or by the mosquito is not clear. The salivary glands of noninfected mosquitoes express low levels of α -gal, as detected by western blot (Figure 1) and immunostaining (data not shown). *Plasmodium* sporozoites are masked by mosquito laminin (Warburg et al., 2007), an evolutionary conserved glycoprotein that in other species contains α -gal (Takahashi et al., 2014). It is possible therefore that anti- α -gal Abs recognize laminin or another mosquito-derived protein expressing α -gal, masking *Plasmodium* sporozoites (Warburg et al., 2007).

It is now well established that specific components of the gut microbiota can modulate immunity and resistance to infection (Belkaid and Hand, 2014; Honda and Littman, 2012). In support of this notion, resistance to viral and bacterial (Fagundes et al., 2012) infections is impaired in GF mice (Dolowy and Muldoon, 1964) or mice subjected to antibiotic-driven dysbiosis (Ichinohe et al., 2011). We reasoned that xeno-glycans expressed by specific components of the gut microbiota might trigger a protective immune response against pathogens expressing the same xeno-glycans. We show that this is the case for α -gal, a xeno-glycan expressed by the human gut pathobiont *E. coli* O86:B7, as well by *Plasmodium* spp. (Figures 1, 3A, and 3B). When colonized by *E. coli* O86:B7, $\alpha 1,3Gt^{-/-}$ mice produce an anti- α -gal IgM Ab response (Figures 3C and 3F) that confer protection against *Plasmodium* infection (Figures 3D, 3E, 3G, and 5C) via a lytic mechanism mediated by complement activation (Figures 5E and 6A). It is worth noticing that in a similar manner to other microbiota-driven resistance mechanisms, the protective effect exerted by *E. coli* O86:B7 acts at the level of a tissue barrier, i.e., the skin, to prevent *Plasmodium* transmission (Figures 4E and 4F).

Levels of circulating anti- α -gal IgG Abs in individuals from a malaria endemic region (Figure 2C), as well as in $\alpha 1,3Gt^{-/-}$ mice colonized with *E. coli* O86:B7 (Figures S2A and S2C), are ~ 30 -fold and ~ 40 to 70-fold, respectively, lower than levels of IgM anti- α -gal Abs. This may explain why basal levels anti- α -gal IgG Abs in individuals from a malaria endemic region are not associated with decreased risk of *P. falciparum* infection (Figure 2D). This also suggests that *P. falciparum* infection fails to induce class switch of the anti- α -gal Ig Ab response in those individuals. It is possible therefore that *P. falciparum* represses Ig class-switch recombination, explaining the residual levels of circulating anti- α -gal IgG Abs (Figure 2C).

While anti- α -gal Abs can provide sterile protection against malaria in mice (Figures 5B, 5C, and S4B), this is not commonly observed in malaria endemic regions in which adult individuals have circulating anti- α -gal IgM Abs, possibly because the levels of these Abs are below a threshold level required to provide sterile protection (Figure 2). However, we show that this natural mechanism of protection can be enhanced via immunization using adjuvants that favor the production of T cell-dependent complement fixing anti- α -gal IgG Abs (Figures 4G and 4H). Moreover, when coupled to *Plasmodium* antigens, this approach should enhance the immunogenicity of such antigens (Benatui et al., 2005) and boost the protective efficacy of candidate malaria vaccines based on such antigens (Olotu et al., 2013). This approach should be useful in preventing not only individual infections but also disease transmission given the protective effect of anti- α -gal Abs.

It is possible that the protective effect of “attenuated” sporozoite vaccine trials against malaria (Seder et al., 2013) is driven to some extent by an anti- α -gal Ab response, given the expression of α -gal by *Plasmodium falciparum* sporozoites (Figure 1). Whether a correlation can be established between the effectiveness of such candidate vaccines and a putative anti- α -gal IgG Ab response has not been established but may be useful to consider as a retrospective analyses.

As a final note, we predict that in a similar manner to anti- α -gal Abs other anti-glycan Abs may confer protection against malaria

as well as other vector-borne protozoan parasites (Huffeijt et al., 2009; Lacroix-Desmazes et al., 1995; Nagele et al., 2013). Moreover, anti- α -gal Abs may also target other vector-borne protozoan parasites expressing α -gal, such as *Leishmania* spp. and *Trypanosoma* spp., the causative agents of Leishmaniasis and Trypanosomiasis, respectively (Avila et al., 1989). As such, vaccination approaches similar to the one proposed here for malaria may be considered against these diseases as well.

EXPERIMENTAL PROCEDURES

Cohort Study

For detailed analysis, see the Extended Experimental Procedures.

Immunization against α -Gal

Eight- to ten-week-old mice received 3×10^8 rabbit rBCM equivalents (100 μ l; PBS; intraperitoneal [i.p.]). Adjuvants are described in the Extended Experimental Procedures. Mouse serum was collected 2 weeks after last immunization, and circulating anti- α -gal Abs were quantified by ELISA. See the Extended Experimental Procedures for details on anti- α -gal ELISA.

Passive Transfer of Anti- α -Gal mAbs

$\alpha 1,3Gt^{-/-}$ mice received anti- α -gal IgG1, IgG2a, IgG2b, and IgG3 mAbs (Ding et al., 2008; Yin et al., 2004) (150 μ g; 100 μ l per mouse) or polyclonal IgM (150 μ g; 300–400 μ l per mouse) via a single intravenous (i.v.) injection 24 hr prior to mosquito exposure.

Plasmodium Strains

Transgenic *P. berghei* ANKA (*PbA*) strains expressing GFP under the *eef1 α* promoter, i.e., *PbA^{Eef1 α -GFP}* (259cl1; MR4; MRA-865) (Franke-Fayard et al., 2004), or under the *hsp70* promoter (Ishino et al., 2006), i.e., *PbA^{Hsp70-GFP}* (kindly provided by Robert Menard, Institut Pasteur), transgenic *P. yoelii* 17XNL strain expressing GFP under the *PbA eef1 α* promoter (MR4; MRA-817; kindly provided by Robert Menard, Institut Pasteur) (Weiss et al., 1989). For sporozoite production, see the Extended Experimental Procedures.

Plasmodium Transmission

A. stephensi or *gambiae* mosquitoes were allowed to feed on anesthetized mice (i.p.; 125 mg/kg ketamine; 12.5 mg/kg xylazine) placed on a warming tray. Two mosquitoes were allowed to probe and feed independently (90–100 s) on restricted to the edge of the mouse ear (10–12/3–4 mm) and dissected thereafter for confirmation of sporozoites in salivary glands. If negative, infection was repeated.

Sporozoites Inoculation

PbA^{Eef1 α -GFP} sporozoites were inoculated (i.d.) in the ear pinna (750 sporozoites in 20–30 μ l; 1% BSA in PBS) or i.v. in the retro-orbital vein (150 sporozoites in 50 μ l; 1% BSA in PBS) using a microsyringe (Nanofil 100 μ l; 33G beveled needle; World Precision Instruments).

Detection of α -Gal in Plasmodium Sporozoites

Sporozoites were stained with Alexa Fluor 647-conjugated BSI-IB₄ or Alexa Fluor 647-conjugated anti- α -gal mAbs and detected by confocal microscope or flow cytometer. For detection of α -gal in *PbA^{Hsp70-GFP}* by western blotting, see the Extended Experimental Procedures. Green coffee bean α -galactosidase (50–200 μ l; 5 U/ml; 60–90 min; 25°C; Sigma Chemical) was used to hydrolyze terminal α -galactosyl moieties from glycolipids and glycoproteins (Luo et al., 1999).

Statistical Analysis

All tests (except human cohort studies) were performed using the GraphPad Prism (v. 6.0) (GraphPad Software). Human analyses were performed in R (v. 3.0.2). Detailed analyses are described in the Extended Experimental Procedures.

Mice

Experiments in mice were performed in accordance with protocols approved by the Ethics Committee of the Instituto Gulbenkian de Ciência and the Portuguese National Entity (DGAV-Direção Geral de Alimentação e Veterinária). Experiments in mice were performed in accordance with the Portuguese (Decreto-Lei no. 113/2013) and European (directive 2010/63/EU) legislation related to housing, husbandry, and animal welfare. C57BL/6 $J_HT^{-/-}$ (Gu et al., 1993), $Tcr\beta^{-/-}$ (Mombaerts et al., 1992), $Aid^{-/-}$ (Muramatsu et al., 2000), $\mu S^{-/-}$ (Ehrenstein et al., 1998), and $C3^{-/-}$ (Circolo et al., 1999) mice were crossed with C57BL/6 $\alpha 1,3Gt^{-/-}$ mice (Shinkel et al., 1997; Tearle et al., 1996). For the details on genotyping, see the Extended Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures and seven figures and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2014.10.053>.

AUTHOR CONTRIBUTIONS

B.Y. contributed to study design, performed, and/or contributed critically to all experiments, analyzed data, and contributed to writing of the manuscript. In some experiments, B.Y. was assisted by S.R., S.P., T.M.T., and P.D.C. designed, performed, and analyzed the human studies. R.G. performed the western blot experiments. H.S. supervised J.G. in the establishment and maintenance of *Plasmodium*-infected *A. mosquitoes*. A.R. produced and trouble-shooted all mAb production. P.J.C. and A.J.F.A. generated $\alpha 1,3Gt^{-/-}$ mice. A.S.C. provided anti- α -gal hybridomas. O.K.D. and B.T. organized the human studies and provided the human serum samples. M.P.S. formulated the original hypothesis, drove the study design, and wrote the manuscript with B.Y. All authors read and approved the manuscript.

ACKNOWLEDGMENTS

The authors thank the Inflammation Group (IGC) for insightful discussions and review of the manuscript, Sofia Rebelo and Sílvia Cardoso for mouse breeding and genotyping, Pedro Almada and Nuno Pimpão Martins (IGC Imaging Facility) for technical support, Karen Berman de Ruiz and Joana Born (IGC Animal Facility) for germ-free breeding, Joana Tavares, Rogerio Amino, and Robert Ménard (Institute Pasteur) for technical support, Alekos Athanasiadis and Jocelyne Demengeot for insightful discussions, Pascal Gagneux (University of California San Diego), and Daniel Mucida (Rockefeller University) for critical review of the initial version of the manuscript. Financial support from the Bill and Melinda Gates Foundation (OPP1024563), Fundação para a Ciência e Tecnologia (RECI-IMI-IMU-0038-2012), and European Research Council (ERC-2011-AdG 294709-DAMAGECONTROL) (to M.P.S.) and Fundação para a Ciência e a Tecnologia (SFRH/BD/51176/2010) within the PhD Program of Integrative Biomedical Science of the Instituto Gulbenkian de Ciência (to B.Y.) is gratefully acknowledged. The Division of Intramural Research, National Institute of Allergy and Infectious Diseases, and NIH supported the Mali cohort study. Mouse axenization was supported by the EMMA, EU FP7 Capacities Specific Program.

Received: August 8, 2014

Revised: September 26, 2014

Accepted: September 30, 2014

Published: December 4, 2014

REFERENCES

Avila, J.L., Rojas, M., and Galili, U. (1989). Immunogenic Gal alpha 1—3Gal carbohydrate epitopes are present on pathogenic American Trypanosoma and Leishmania. *J. Immunol.* 142, 2828–2834.

Avila, J.L., Rojas, M., and Velazquez-Avila, G. (1992). Characterization of a natural human antibody with anti-galactosyl(alpha 1-2)galactose specificity that is

present at high titers in chronic Trypanosoma cruzi infection. *Am. J. Trop. Med. Hyg.* 47, 413–421.

Belkaid, Y., and Hand, T.W. (2014). Role of the microbiota in immunity and inflammation. *Cell* 157, 121–141.

Belnoue, E., Kayibanda, M., Vigario, A.M., Deschemin, J.C., van Rooijen, N., Viguier, M., Snounou, G., and Rénia, L. (2002). On the pathogenic role of brain-sequestered alphabeta CD8+ T cells in experimental cerebral malaria. *J. Immunol.* 169, 6369–6375.

Benatui, L., Kaye, J., Rich, R.F., Fishman, J.A., Green, W.R., and Iacomini, J. (2005). The influence of natural antibody specificity on antigen immunogenicity. *Eur. J. Immunol.* 35, 2638–2647.

Bishop, J.R., and Gagneux, P. (2007). Evolution of carbohydrate antigens—microbial forces shaping host glycomes? *Glycobiology* 17, 23R–34R.

Chen, L., Zhang, Z., and Sando, F. (2000). Neutrophils play a critical role in the pathogenesis of experimental cerebral malaria. *Clin. Exp. Immunol.* 120, 125–133.

Chiang, T.R., Fanget, L., Gregory, R., Tang, Y., Ardiet, D.L., Gao, L., Meschter, C., Kozikowski, A.P., Buelow, R., and Vuist, W.M. (2000). Anti-gal antibodies in humans and 1, 3alpha-galactosyltransferase knock-out mice. *Transplantation* 69, 2593–2600.

Cretin, N., Bracy, J., Hanson, K., and Iacomini, J. (2002). The role of T cell help in the production of antibodies specific for Gal alpha 1-3Gal. *J. Immunol.* 168, 1479–1483.

Cywes-Bentley, C., Skurnik, D., Zaidi, T., Roux, D., Deoliveira, R.B., Garrett, W.S., Lu, X., O'Malley, J., Kinzel, K., Zaidi, T., et al. (2013). Antibody to a conserved antigenic target is protective against diverse prokaryotic and eukaryotic pathogens. *Proc. Natl. Acad. Sci. USA* 110, E2209–E2218.

Ding, J.W., Zhou, T., Zeng, H., Ma, L., Verbeek, J.S., Yin, D., Shen, J., and Chong, A.S. (2008). Hyperacute rejection by anti-Gal IgG1, IgG2a, and IgG2b is dependent on complement and Fc-gamma receptors. *J. Immunol.* 180, 261–268.

Doenz, U., Nydegger, U.E., Kueng, A., Carrel, T., and Mohacs, P. (2000). Anti-Galalpha1-3Gal IgM/IgG antibody levels in infants: do they have a clinical relevance in pediatric xenotransplantation? *J. Heart Lung Transplant.* 19, 1108–1113.

Dolowy, W.C., and Muldoon, R.L. (1964). Studies of germfree animals. I. Response of mice to infection with influenza A virus. *Proc. Soc. Exp. Biol. Med.* 116, 365–371.

Ehrenstein, M.R., O'Keefe, T.L., Davies, S.L., and Neuberger, M.S. (1998). Targeted gene disruption reveals a role for natural secretory IgM in the maturation of the primary immune response. *Proc. Natl. Acad. Sci. USA* 95, 10089–10093.

Fagundes, C.T., Amaral, F.A., Vieira, A.T., Soares, A.C., Pinho, V., Nicoli, J.R., Vieira, L.Q., Teixeira, M.M., and Souza, D.G. (2012). Transient TLR activation restores inflammatory response and ability to control pulmonary bacterial infection in germfree mice. *J. Immunol.* 188, 1411–1420.

Franke-Fayard, B., Trueman, H., Ramesar, J., Mendoza, J., van der Keur, M., van der Linden, R., Sinden, R.E., Waters, A.P., and Janse, C.J. (2004). A Plasmodium berghei reference line that constitutively expresses GFP at a high level throughout the complete life cycle. *Mol. Biochem. Parasitol.* 137, 23–33.

Galili, U., and Swanson, K. (1991). Gene sequences suggest inactivation of alpha-1,3-galactosyltransferase in catarrhines after the divergence of apes from monkeys. *Proc. Natl. Acad. Sci. USA* 88, 7401–7404.

Galili, U., LaTemple, D.C., and Radic, M.Z. (1998). A sensitive assay for measuring alpha-Gal epitope expression on cells by a monoclonal anti-Gal antibody. *Transplantation* 65, 1129–1132.

Galili, U., Rachmilewitz, E.A., Peleg, A., and Flechner, I. (1984). A unique natural human IgG antibody with anti-alpha-galactosyl specificity. *J. Exp. Med.* 160, 1519–1531.

Galili, U., Macher, B.A., Buehler, J., and Shohet, S.B. (1985). Human natural anti-alpha-galactosyl IgG. II. The specific recognition of alpha (1—3)-linked galactose residues. *J. Exp. Med.* 162, 573–582.

- Galili, U., Mandrell, R.E., Hamadeh, R.M., Shohet, S.B., and Griffiss, J.M. (1988). Interaction between human natural anti- α -galactosyl immunoglobulin G and bacteria of the human flora. *Infect. Immun.* 56, 1730–1737.
- Gu, H., Zou, Y.R., and Rajewsky, K. (1993). Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. *Cell* 73, 1155–1164.
- Hayakawa, T., Satta, Y., Gagneux, P., Varki, A., and Takahata, N. (2001). Alu-mediated inactivation of the human CMP- N-acetylneuraminic acid hydroxylase gene. *Proc. Natl. Acad. Sci. USA* 98, 11399–11404.
- Hegge, S., Kudryashev, M., Barniol, L., and Frischknecht, F. (2010). Key factors regulating *Plasmodium berghei* sporozoite survival and transformation revealed by an automated visual assay. *FASEB journal* 24, 5003–5012.
- Honda, K., and Littman, D.R. (2012). The microbiome in infectious disease and inflammation. *Annu. Rev. Immunol.* 30, 759–795.
- Hufejt, M.E., Vuskovic, M., Vasiliu, D., Xu, H., Obukhova, P., Shilova, N., Tuzikov, A., Galanina, O., Arun, B., Lu, K., and Bovin, N. (2009). Anti-carbohydrate antibodies of normal sera: findings, surprises and challenges. *Mol. Immunol.* 46, 3037–3049.
- Ichinohe, T., Pang, I.K., Kumamoto, Y., Peaper, D.R., Ho, J.H., Murray, T.S., and Iwasaki, A. (2011). Microbiota regulates immune defense against respiratory tract influenza A virus infection. *Proc. Natl. Acad. Sci. USA* 108, 5354–5359.
- Ishino, T., Orito, Y., Chinzei, Y., and Yuda, M. (2006). A calcium-dependent protein kinase regulates *Plasmodium* ookinete access to the midgut epithelial cell. *Mol. Microbiol.* 59, 1175–1184.
- Kimmel, J., Ogun, S.A., de Macedo, C.S., Gerold, P., Vivas, L., Holder, A.A., Schwarz, R.T., and Azzouz, N. (2003). Glycosylphosphatidylinositol in murine malaria: *Plasmodium yoelii yoelii*. *Biochimie* 85, 473–481.
- Lacroix-Desmazes, S., Mouthon, L., Coutinho, A., and Kazatchkine, M.D. (1995). Analysis of the natural human IgG antibody repertoire: life-long stability of reactivities towards self antigens contrasts with age-dependent diversification of reactivities against bacterial antigens. *Eur. J. Immunol.* 25, 2598–2604.
- Luo, Y., Wen, J., Luo, C., Cummings, R.D., and Cooper, D.K. (1999). Pig xenogeneic antigen modification with green coffee bean α -galactosidase. *Xenotransplantation* 6, 238–248.
- Macher, B.A., and Galili, U. (2008). The Gal α 1,3Gal β 1,4GlcNAc-R (α -Gal) epitope: a carbohydrate of unique evolution and clinical relevance. *Biochim. Biophys. Acta* 1780, 75–88.
- Mañez, R., Blanco, F.J., Díaz, I., Centeno, A., Lopez-Pelaez, E., Hermida, M., Davies, H.F., and Katopodis, A. (2001). Removal of bowel aerobic gram-negative bacteria is more effective than immunosuppression with cyclophosphamide and steroids to decrease natural α -galactosyl IgG antibodies. *Xenotransplantation* 8, 15–23.
- Ménard, R., Tavares, J., Cockburn, I., Markus, M., Zavala, F., and Amino, R. (2013). Looking under the skin: the first steps in malarial infection and immunity. *Nat. Rev. Microbiol.* 11, 701–712.
- Miyatake, T., Sato, K., Takigami, K., Koyamada, N., Hancock, W.W., Bazin, H., Latinne, D., Bach, F.H., and Soares, M.P. (1998). Complement-fixing elicited antibodies are a major component in the pathogenesis of xenograft rejection. *J. Immunol.* 160, 4114–4123.
- Modiano, D., Petrarca, V., Sirima, B.S., Nebié, I., Diallo, D., Esposito, F., and Coluzzi, M. (1996). Different response to *Plasmodium falciparum* malaria in west African sympatric ethnic groups. *Proc. Natl. Acad. Sci. USA* 93, 13206–13211.
- Mombaerts, P., Clarke, A.R., Rudnicki, M.A., Iacomini, J., Itohara, S., Lafaille, J.J., Wang, L., Ichikawa, Y., Jaenisch, R., Hooper, M.L., et al. (1992). Mutations in T-cell antigen receptor genes α and β block thymocyte development at different stages. *Nature* 360, 225–231.
- Moran, P., and Caras, I.W. (1994). Requirements for glycosylphosphatidylinositol attachment are similar but not identical in mammalian cells and parasitic protozoa. *J. Cell Biol.* 125, 333–343.
- Mota, M.M., Pradel, G., Vanderberg, J.P., Hafalla, J.C., Frevert, U., Nussenzweig, R.S., Nussenzweig, V., and Rodríguez, A. (2001). Migration of *Plasmodium* sporozoites through cells before infection. *Science* 291, 141–144.
- Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., and Honjo, T. (2000). Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 102, 553–563.
- Nagele, E.P., Han, M., Acharya, N.K., DeMarshall, C., Kosciuk, M.C., and Nagele, R.G. (2013). Natural IgG autoantibodies are abundant and ubiquitous in human sera, and their number is influenced by age, gender, and disease. *PLoS ONE* 8, e60726.
- Nimmerjahn, F., and Ravetch, J.V. (2008). Fc γ receptors as regulators of immune responses. *Nat. Rev. Immunol.* 8, 34–47.
- Olotu, A., Fegan, G., Wambua, J., Nyangweso, G., Awuondo, K.O., Leach, A., Lievens, M., Lebouilleux, D., Njuguna, P., Peshu, N., et al. (2013). Four-year efficacy of RTS,S/AS01E and its interaction with malaria exposure. *N. Engl. J. Med.* 368, 1111–1120.
- Oyelaran, O., McShane, L.M., Dodd, L., and Gildersleeve, J.C. (2009). Profiling human serum antibodies with a carbohydrate antigen microarray. *J. Proteome Res.* 8, 4301–4310.
- Pal, S.C., Rao, C.K., Kereselidze, T., Krishnaswami, A.K., Murty, D.K., Pandit, C.G., and Shrivastav, J.B. (1969). An extensive community outbreak of enteropathogenic *Escherichia coli* O86: B7 gastroenteritis. *Bull. World Health Organ.* 41, 851–858.
- Parker, W., Lin, S.S., Yu, P.B., Sood, A., Nakamura, Y.C., Song, A., Everett, M.L., and Platt, J.L. (1999). Naturally occurring anti- α -galactosyl antibodies: relationship to xenoreactive anti- α -galactosyl antibodies. *Glycobiology* 9, 865–873.
- Porcherie, A., Mathieu, C., Peronet, R., Schneider, E., Claver, J., Commere, P.H., Kiefer-Biasizzo, H., Karasuyama, H., Milon, G., Dy, M., et al. (2011). Critical role of the neutrophil-associated high-affinity receptor for IgE in the pathogenesis of experimental cerebral malaria. *J. Exp. Med.* 208, 2225–2236.
- Posekany, K.J., Pittman, H.K., Bradfield, J.F., Haisch, C.E., and Verbanac, K.M. (2002). Induction of cytolytic anti-Gal antibodies in α -1,3-galactosyltransferase gene knockout mice by oral inoculation with *Escherichia coli* O86:B7 bacteria. *Infect. Immun.* 70, 6215–6222.
- Rickman, L.S., Jones, T.R., Long, G.W., Paparello, S., Schneider, I., Paul, C.F., Beaudoin, R.L., and Hoffman, S.L. (1990). *Plasmodium falciparum*-infected *Anopheles stephensi* inconsistently transmit malaria to humans. *Am. J. Trop. Med. Hyg.* 43, 441–445.
- Ringel-Kulka, T., Cheng, J., Ringel, Y., Salojärvi, J., Carroll, I., Palva, A., de Vos, W.M., and Satokari, R. (2013). Intestinal microbiota in healthy U.S. young children and adults—a high throughput microarray analysis. *PLoS ONE* 8, e64315.
- Sauerwein, R.W., Roestenberg, M., and Moorthy, V.S. (2011). Experimental human challenge infections can accelerate clinical malaria vaccine development. *Nat. Rev. Immunol.* 11, 57–64.
- Seder, R.A., Chang, L.J., Enama, M.E., Zephir, K.L., Sarwar, U.N., Gordon, I.J., Holman, L.A., James, E.R., Billingsley, P.F., Gunasekera, A., et al.; VRC 312 Study Team (2013). Protection against malaria by intravenous immunization with a nonreplicating sporozoite vaccine. *Science* 341, 1359–1365.
- Springer, G.F., and Horton, R.E. (1969). Blood group isoantibody stimulation in man by feeding blood group-active bacteria. *J. Clin. Invest.* 48, 1280–1291.
- Springer, G.F., Horton, R.E., and Forbes, M. (1959). [Origin of anti-human blood group B agglutinins in white Leghorn chicks]. *J. Exp. Med.* 110, 221–244.
- Takahashi, H., Chinuki, Y., Tanaka, A., and Morita, E. (2014). Laminin γ -1 and collagen α -1 (VI) chain are galactose- α -1,3-galactose-bound allergens in beef. *Allergy* 69, 199–207.
- Takeuchi, Y., Porter, C.D., Strahan, K.M., Preece, A.F., Gustafsson, K., Cosset, F.L., Weiss, R.A., and Collins, M.K. (1996). Sensitization of cells and retroviruses to human serum by (α -1-3) galactosyltransferase. *Nature* 379, 85–88.

- Tangvoranuntakul, P., Gagneux, P., Diaz, S., Bardor, M., Varki, N., Varki, A., and Muchmore, E. (2003). Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *Proc. Natl. Acad. Sci. USA* *100*, 12045–12050.
- Tearle, R.G., Tange, M.J., Zannettino, Z.L., Katerelos, M., Shinkel, T.A., Van Denderen, B.J., Lonie, A.J., Lyons, I., Nottle, M.B., Cox, T., et al. (1996). The alpha-1,3-galactosyltransferase knockout mouse. Implications for xenotransplantation. *Transplantation* *61*, 13–19.
- Tran, T.M., Ongoiba, A., Coursen, J., Crosnier, C., Diouf, A., Huang, C.Y., Li, S., Doumbo, S., Doumtabe, D., Kone, Y., et al. (2014). Naturally acquired antibodies specific for *Plasmodium falciparum* reticulocyte-binding protein homologue 5 inhibit parasite growth and predict protection from malaria. *J. Infect. Dis.* *209*, 789–798.
- Verhage, D.F., Telgt, D.S., Bousema, J.T., Hermesen, C.C., van Gemert, G.J., van der Meer, J.W., and Sauerwein, R.W. (2005). Clinical outcome of experimental human malaria induced by *Plasmodium falciparum*-infected mosquitoes. *Neth. J. Med.* *63*, 52–58.
- Warburg, A., Shtern, A., Cohen, N., and Dahan, N. (2007). Laminin and a *Plasmodium* ookinete surface protein inhibit melanotic encapsulation of Sephadex beads in the hemocoel of mosquitoes. *Microbes Infect.* *9*, 192–199.
- Weiss, W.R., Good, M.F., Hollingdale, M.R., Miller, L.H., and Berzofsky, J.A. (1989). Genetic control of immunity to *Plasmodium yoelii* sporozoites. *J. Immunol.* *143*, 4263–4266.
- Yang, Y.G., deGoma, E., Ohdan, H., Bracy, J.L., Xu, Y., Iacomini, J., Thall, A.D., and Sykes, M. (1998). Tolerization of anti-Galalpha1-3Gal natural antibody-forming B cells by induction of mixed chimerism. *J. Exp. Med.* *187*, 1335–1342.
- Yi, W., Bystricky, P., Yao, Q., Guo, H., Zhu, L., Li, H., Shen, J., Li, M., Ganguly, S., Bush, C.A., and Wang, P.G. (2006). Two different O-polysaccharides from *Escherichia coli* O86 are produced by different polymerization of the same O-repeating unit. *Carbohydr. Res.* *341*, 100–108.
- Yin, D., Zeng, H., Ma, L., Shen, J., Xu, H., Byrne, G.W., and Chong, A.S. (2004). Cutting edge: NK cells mediate IgG1-dependent hyperacute rejection of xenografts. *J. Immunol.* *172*, 7235–7238.

Host Adaptation of a Bacterial Toxin from the Human Pathogen *Salmonella* Typhi

Lingquan Deng,^{1,2,4,8} Jeongmin Song,^{5,8,9} Xiang Gao,^{5,8} Jiawei Wang,⁶ Hai Yu,⁷ Xi Chen,⁷ Nissi Varki,^{1,3} Yuko Naito-Matsui,^{1,2,4} Jorge E. Galán,^{5,*} and Ajit Varki^{1,2,3,4,*}

¹Glycobiology Research and Training Center

²Department of Medicine

³Department of Pathology

⁴Department of Cellular and Molecular Medicine

University of California, San Diego, La Jolla, CA 92093, USA

⁵Department of Microbial Pathogenesis, Yale University School of Medicine, New Haven, CT 06536, USA

⁶School of Life Sciences, Tsinghua University, Beijing 100084, PRC

⁷Department of Chemistry, University of California, Davis, CA 95616, USA

⁸Co-first author

⁹Present address: Department of Microbiology and Immunology, Cornell University College of Veterinary Medicine, Ithaca, NY 14853, USA

*Correspondence: jorge.galan@yale.edu (J.E.G.), a1varki@ucsd.edu (A.V.)

<http://dx.doi.org/10.1016/j.cell.2014.10.057>

SUMMARY

Salmonella Typhi is an exclusive human pathogen that causes typhoid fever. Typhoid toxin is a *S. Typhi* virulence factor that can reproduce most of the typhoid fever symptoms in experimental animals. Toxicity depends on toxin binding to terminally sialylated glycans on surface glycoproteins. Human glycans are unusual because of the lack of CMAH, which in other mammals converts N-acetylneuraminic acid (Neu5Ac) to N-glycolylneuraminic acid (Neu5Gc). Here, we report that typhoid toxin binds to and is toxic toward cells expressing glycans terminated in Neu5Ac (expressed by humans) over glycans terminated in Neu5Gc (expressed by other mammals). Mice constitutively expressing CMAH thus displaying Neu5Gc in all tissues are resistant to typhoid toxin. The atomic structure of typhoid toxin bound to Neu5Ac reveals the structural bases for its binding specificity. These findings provide insight into the molecular bases for *Salmonella* Typhi's host specificity and may help the development of therapies for typhoid fever.

INTRODUCTION

Salmonella enterica serovar Typhi (*S. Typhi*), the cause of typhoid fever, continues to be a major public health concern, particularly in developing countries. There are more than 20 million cases of typhoid fever every year, which result in more than 400,000 deaths (Crump and Mintz, 2010; Parry and Threlfall, 2008; Voetsch et al., 2004). Unlike the illnesses associated with most other *Salmonella*, which are usually self-limiting gastroenteritis (i.e., “food poisoning”), typhoid fever is a systemic, often lethal disease (Parry et al., 2002). In addition, in contrast to most *Salmonella enterica* serovars, which can infect

a broad range of hosts, *S. Typhi* exhibits remarkable host specificity, causing symptomatic infections only in humans (Ohi and Miller, 2001; Parry et al., 2002; Raffatellu et al., 2008). The mechanisms of *S. Typhi* host specificity are incompletely understood and most likely multifactorial. For example, *S. Typhi* is unable to replicate in most hosts, except chimpanzees where it was found to reach levels equivalent to those in humans (Edsall et al., 1960; Metchnikoff and Besredka, 1911). However, despite significant bacterial replication, *S. Typhi* did not cause typical typhoid fever symptoms in chimpanzees, which developed a milder and much shorter lasting disease syndrome (Edsall et al., 1960; Metchnikoff and Besredka, 1911). These observations indicate that in addition to pathogen restriction, there are other host factors that must prevent the development of typhoid fever even in the presence of significant bacterial replication. Host restriction is manifested at the cellular level because, in contrast to human macrophages, *S. Typhi* is unable to survive within macrophages of nonpermissive species (Schwan et al., 2000; Vladoianu et al., 1990). Recent studies have identified a Rab32-dependent pathogen-restriction mechanism that limits the growth of *S. Typhi* within macrophages of nonpermissive species (Spanò and Galán, 2012). In contrast, this antimicrobial function is effectively neutralized by broad-host *Salmonella* serovars, which are able to proteolytically target Rab32 with a type III secretion effector protein that is absent from *S. Typhi* (Spanò and Galán, 2012; Spanò et al., 2011).

Typhoidal (i.e., able to cause typhoid fever) *Salmonella* serovars (e.g., *S. Typhi* and *S. Paratyphi*) encode typhoid toxin, a unique member of the AB₅ exotoxin family (Haghjoo and Galán, 2004; Song et al., 2013; Spanò et al., 2008). Unlike all known members of this family, which possess a single enzymatic A subunit associated to a pentameric B subunit (Beddoe et al., 2010), typhoid toxin is composed of two covalently linked enzymatic subunits, the deoxyribonuclease CdtB and the ADP ribosyl transferase PliA, associated to the homopentameric B subunit PliB (Song et al., 2013). Thus typhoid toxin may have evolved from the combination of two exotoxins, cytolethal distending and pertussis toxins, and is the only known example of a toxin

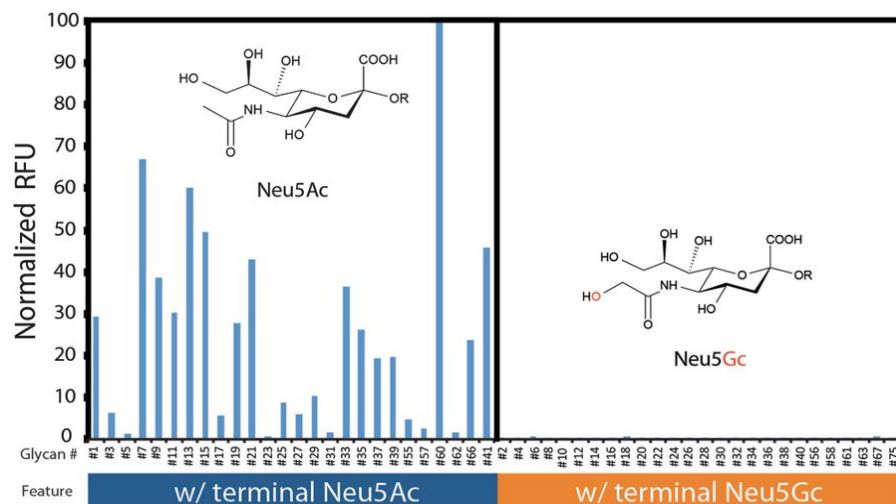


Figure 1. Comparison of Typhoid Toxin Binding to Paired Neu5Ac- and Neu5Gc-Terminated Glycans by a Customized Microarray

Chemical structures of Neu5Ac and Neu5Gc are shown. The two molecules differ by only one single oxygen atom. Vertical axis values represent the normalized average of relative fluorescence units and horizontal axis indicates the glycan number in the array.

See also Figure S1 and Tables 1 and S1.

with an A₂B₅ organization. Recent studies have shown that direct injection of typhoid toxin into experimental animals can reproduce many of the pathognomonic symptoms of typhoid fever, thus placing this toxin at the center of the pathogenesis of this devastating disease (Song et al., 2013).

To enter cells typhoid toxin must bind glycosylated surface glycoprotein receptors in target cells, such as podocalyxin 1 on epithelial cells and CD45 on myelocytic cells (Song et al., 2013). The toxin recognizes specific sialylated glycan moieties on the receptor proteins through a glycan-binding domain in its P1tB B subunit. Sialoglycans on human cells are unusual in that they are primarily terminated in *N*-acetylneuraminic acid (Neu5Ac) (Varki et al., 2011). This is in contrast to other old world primates and most other mammals studied to date, whose glycans can also terminate in *N*-glycolylneuraminic acid (Neu5Gc). These differences in glycan composition are the result of the absence of CMP-*N*-acetylneuraminic acid hydroxylase (CMAH) in humans, due to an Alu-mediated exon deletion in the *CMAH* gene, which occurred after the separation of the *Hominin* lineage from other Hominids (the so-called “great apes,” e.g., chimpanzees) (Chou et al., 2002). Here, we report that typhoid toxin exhibits exquisite specificity for human-like Neu5Ac-terminated glycans. We find that typhoid toxin is cytotoxic to cells expressing Neu5Ac glycans on their surface but not to those expressing Neu5Gc. Furthermore, typhoid toxin binds strongly to human tissues but poorly to those from chimpanzees, which predominantly display Neu5Gc-terminated glycans and do not develop the typical symptoms of typhoid fever. We also show that mice engineered to display Neu5Gc glycans in all tissues are resistant to typhoid toxin. These findings provide major insight into the molecular bases for the host specificity of *S. Typhi* and may help the development of novel therapeutic approaches against typhoid fever.

RESULTS

Typhoid Toxin Exhibits Strong Specificity for Neu5Ac-Terminated Glycans

Given the remarkable human specificity exhibited by *S. Typhi* and the central role of typhoid toxin in the pathogenesis of

typhoid fever, we used a customized glycan array (Padler-Karavani et al., 2014) to compare the ability of fluorescently labeled typhoid toxin to bind pairs of sialylated glycans terminated in either Neu5Ac (predominantly expressed in human cells) or Neu5Gc (predominantly expressed in cells of most other mammals). Consistent with previous results (Song et al., 2013), typhoid toxin bound a diverse group of sialylated glycans with preferential binding to termini with the consensus sequence Neu5Ac α 2-3Gal β 1-3/ β 1-4Glc/GlcNAc (Figure 1; Table 1; Table S1 available online). Remarkably, however, we found that typhoid toxin did not bind to otherwise identical glycans terminated in Neu5Gc (i.e., differing by a single oxygen atom) (Figures 1 and S1; Table 1; Table S1). The marked difference in binding was observed across all the glycans tested and with different toxin concentrations (Table 1; Figure S1; Table S1). Consistent with previous observations (Song et al., 2013), typhoid toxin carrying a mutation in the glycan-binding site of its P1tB B subunit (P1tB^{S35A}) did not show significant binding to any of the glycans tested regardless of their sialylation status (Table S1). To confirm typhoid toxin's preference for Neu5Ac-terminated glycans, we compared its binding to human and chimpanzee red blood cells, which display markedly different levels of surface Neu5Ac or Neu5Gc-terminated glycans (Figure 2A). In keeping with the glycan array findings, typhoid toxin showed much stronger binding to human than to chimpanzee cells (Figure 2B). Furthermore, the differences were observed at different toxin concentrations (Figure 2B). Similar differences were observed with lymphocytes from humans and chimpanzees (Figures 2C and 2D). These results indicate that typhoid toxin exhibits strong binding preference for Neu5Ac-terminated glycans, which are predominant in human cells.

Incorporation of Neu5Gc Renders Human Cells Resistant to Typhoid Toxin

Although human cells lack Neu5Gc, they can metabolically incorporate this sialic acid if supplied into the cell culture medium (Tangvoranuntakul et al., 2003). To investigate the biological significance of typhoid toxin glycan selectivity, we compared the binding of fluorescently labeled typhoid toxin to human Henle-407 epithelial cells that had been grown in media supplemented with Neu5Gc or Neu5Ac. Predictably, growth in the presence of Neu5Gc significantly altered the sialic acid composition in these cells resulting in up to 60% of the total sialic acid containing

Table 1. Analysis of Fine Ligand Specificity of Native Typhoid Toxin by a Customized Sialoglycan Microarray

Glycan Structure	Mean Relative Fluorescence Units (n = 4)	
	Sia = Neu5Ac	Sia = Neu5Gc
Sia α 3Gal β 3GlcNAc β 3Gal β 4Glc β O(CH ₂) ₃ NH ₂	34,441	25
Sia9Ac α 3Gal β 3GlcNAc β O(CH ₂) ₃ NH ₂	23,120	73
Sia α 3Gal β 3GlcNAc β O(CH ₂) ₃ NH ₂	20,664	156
Sia α 3Gal β 3GalNAc α O(CH ₂) ₃ NH ₂	17,100	88
Sia α 8Sia α 3Gal β 4Glc β O(CH ₂) ₃ NH ₂	15,821	134
Sia α 3Gal β 4Glc β O(CH ₂) ₃ NH ₂	14,781	108
Sia9Ac α 3Gal β 3GalNAc α O(CH ₂) ₃ NH ₂	13,299	42
Sia α 3Gal β 3GalNAc β O(CH ₂) ₃ NH ₂	12,564	147
Sia α 3Gal β 4GlcNAc β O(CH ₂) ₃ NH ₂	10,468	114
Sia9Ac α 3Gal β 4GlcNAc β O(CH ₂) ₃ NH ₂	10,076	30
Sia α 6Gal β 4Glc β O(CH ₂) ₃ NH ₂	9,605	79
Sia9Ac α 3Gal β 3GalNAc β O(CH ₂) ₃ NH ₂	8,993	72
Sia α 3(Neu5Ac α 6)Gal β 4Glc β O(CH ₂) ₃ NH ₂	8,205	246
Sia9Ac α 3Gal β 4Glc β O(CH ₂) ₃ NH ₂	6,730	76
Sia9Ac α 6Gal β 4Glc β O(CH ₂) ₃ NH ₂	6,619	56
Sia9Ac α 3Gal β O(CH ₂) ₃ NH ₂	3,513	152
Sia α 3Gal β O(CH ₂) ₃ NH ₂	3,040	182
Sia9Ac α 6Gal β 4GlcNAc β O(CH ₂) ₃ NH ₂	2,160	89
Sia α 6Gal β O(CH ₂) ₃ NH ₂	2,078	125
Sia α 6Gal β 4GlcNAc β O(CH ₂) ₃ NH ₂	1,979	229
Sia α 3Gal β 4(Fuc α 3)GlcNAc β O(CH ₂) ₃ NH ₂	1,633	51
Sia α 3Gal β 4(Fuc α 3)GlcNAc6S β O(CH ₂) ₃ NH ₂	827	96
Sia9Ac α 6Gal β O(CH ₂) ₃ NH ₂	592	133
Sia α 3Gal β 4GlcNAc6S β O(CH ₂) ₃ NH ₂	511	80
Sia α 6GalNAc α O(CH ₂) ₃ NH ₂	470	200
Sia9Ac α 6GalNAc α O(CH ₂) ₃ NH ₂	217	49

Neu5Gc (Figure 3A). In contrast, cells grown in standard media or supplemented with Neu5Ac had almost undetectable levels of Neu5Gc (Figure 3A). Consistent with its reduced binding affinity to Neu5Gc terminated glycans, typhoid toxin binding and toxicity was markedly reduced in cells grown in the presence of Neu5Gc (Figures 3B–3D and S2). In contrast, levels of toxin binding and cytotoxicity were slightly increased in cells grown in media supplemented with Neu5Ac in comparison to cells grown in standard media (Figures 3B–3D and S2). Equivalent results were observed in Jurkat T cells, which display a different glycoprotein receptor for typhoid toxin (Song et al., 2013) (Figures 3E–3H). These results indicate that the abundance of Neu5Gc on the surface glycans render cells less permissive for toxin binding and therefore resistant to typhoid toxin.

Constitutive Expression of Neu5Gc Renders Mice Resistant to Typhoid Toxin

Although wild-type mice express a fully functional CMAH and can express Neu5Gc, most of their tissues contain a significant amount of Neu5Ac as well, presumably due to low expression

of CMAH in some cells (Hedlund et al., 2007). Consistent with this observation, *Cmah*^{−/−} mice (Hedlund et al., 2007) did not exhibit significantly higher susceptibility to typhoid toxin, showing equivalent levels of weight loss, white blood cell depletion, and time to death compared to those observed in wild-type animals (Figures 4A–4D). In fact, these mice exhibited a slightly reduced susceptibility to the toxin, which is most likely due to the presence of high levels of Neu5Ac-containing glycoproteins in the sera of these animals, which may provide protection against toxicity by competing with toxin binding to tissue receptors (Beddoe et al., 2010). We therefore tested the susceptibility to typhoid toxin of mice engineered to constitutively overexpress CMAH in all tissues (*Cmah*^{tg}) (Figures 4A–4D). We reasoned that the forced expression of CMAH would result in the predominant display of Neu5Gc on the surface of all cells and thus confer protection against intoxication. Consistent with this hypothesis, *Cmah*^{tg} mice were completely resistant to systemic administration of toxin amounts much higher than those that would be lethal for wild-type mice (Figure 4D). Also consistent with these observations, the binding of fluorescently labeled typhoid toxin was undetectable in tissues from *Cmah*^{tg} mice but was readily detectable in tissues from wild-type and *Cmah*^{−/−} mice (Figure S3).

Typhoid Toxin Does Not Bind to Chimpanzee Tissues

Previous studies have shown that chimpanzees can be experimentally infected with *S. Typhi* (Edsall et al., 1960; Metchnikoff and Besredka, 1911). However, infected animals did not develop the typical symptomatology associated with typhoid fever, such as the presence of stupor and extreme lethargy, showing instead a course of disease that was mild and brief. We found that, contrary to human tissues, toxin binding to chimpanzee organ tissue sections was nondetectable (Figure 5). This is explained by the observation that chimpanzee cells predominantly display Neu5Gc-terminated glycans on their surface and are therefore not permissive for toxin binding. These results are consistent with the hypothesis that the lack of pathognomonic typhoid fever symptomatology observed in chimpanzees experimentally infected with *S. Typhi* is due to the inability of typhoid toxin to gain access to target cells.

The Crystal Structure of Typhoid Toxin B Subunit PltB Bound to Its Sialic Acid Ligand Reveals Structural Bases for Its Binding Specificity

To gain insight into the structural bases for typhoid toxin's binding specificity, we determined the atomic structure at 1.92 Å resolution of PltB bound to GalNAc β 1-4(Neu5Ac α 2-8Neu5Ac α 2-3)Gal β 1-4Glc (Figures 6A and 6B; Table S2), which previous studies have shown binds typhoid toxin with high affinity (Song et al., 2013). Clear electron density corresponding to the (Neu5Ac α 2-8Neu5Ac α 2-3)Gal trisaccharide was unambiguously observed at the canonical glycan-binding sites in two of the five subunits of the PltB pentamer (Figures 6A–6C). Because the glycan-binding sites of every subunit are identical, absence of binding to some subunits is most likely due to their limited accessibility because of crystal packing. No specific contacts between the galactose moiety and PltB were observed in the structure (Figure S4). In contrast, the first of the two Neu5Ac moieties interacts through multiple direct hydrogen bonds and

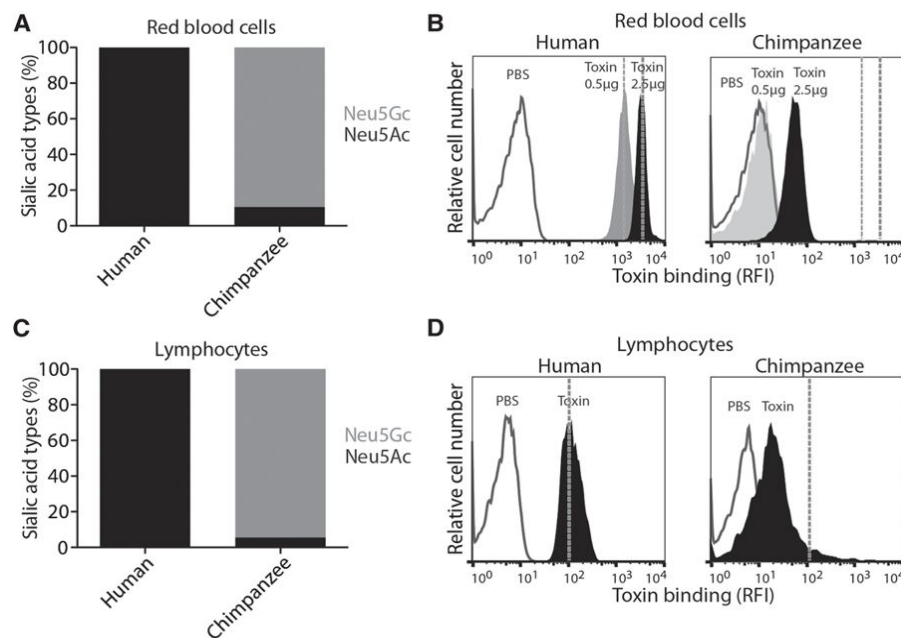


Figure 2. Typhoid Toxin Binding to Red Blood Cells and Lymphocytes from Humans and Chimpanzees

(A) Relative levels of Neu5Ac/Neu5Gc in human and chimpanzee red blood cells (RBCs).

(B) Binding of different amounts of typhoid toxin to human and chimpanzee RBCs.

(C) Relative levels of Neu5Ac/Neu5Gc in human and chimpanzee lymphocytes.

(D) Binding of typhoid toxin to human and chimpanzee lymphocytes. Similar results were obtained in several independent repetitions of the experiments. RFI, relative fluorescence intensity. PBS, phosphate buffered saline.

and suggest an evolutionary pathway by which this toxin restricted its binding to human-specific glycans.

DISCUSSION

Unlike most other *Salmonella enterica* serovars, which can infect a broad range of hosts, *S. Typhi* can only infect humans, in

whom it causes typhoid fever, a severe, often lethal disease. The process by which *S. Typhi* has lost its ability to explore other niches and evolved to cause disease only in humans is incompletely understood and likely to be multifactorial (Jacobsen et al., 2011; Sabbagh et al., 2010). The *S. Typhi* genome sequence exhibits an unusually high number of pseudogenes, suggesting that genome reduction most likely played a central role in its adaptation to a single host (Parkhill et al., 2001). For example, the interaction of *Salmonella enterica* with host cells is largely dictated by two type III secretion systems (T3SS), which deliver bacterial effector proteins into host cells to modulate cellular functions (Galán, 2001; Ibarra and Steele-Mortimer, 2009; Srikanth et al., 2011; Waterman and Holden, 2003). Although these systems are highly conserved across different *Salmonella* serovars, the effectors they deliver are not, and *S. Typhi* expresses a significantly smaller number of effector proteins than most other serovars. One of the missing effectors from *S. Typhi* is GtgE, which is involved in neutralizing a host restriction pathway that prevents its growth in macrophages of nonpermissive species (Spanò et al., 2011).

S. Typhi host specificity, however, is not exclusively due to its inability to replicate within nonpermissive hosts. For example, unlike mice, chimpanzees were found to be permissive for *S. Typhi* replication, and experimental infections showed that in these animals it reached levels equivalent to those observed in infected humans (Edsall et al., 1960; Metchnikoff and Besredka, 1911). However, chimpanzees did not develop the typical symptomatology of typhoid fever indicating that factors other than pathogen restriction contribute to *S. Typhi*'s host specificity. Previous studies have shown that typhoid toxin is central for the development of pathognomonic symptoms of typhoid fever (Song et al., 2013). We have shown here that typhoid toxin exhibits strong selectivity for Neu5Ac-terminated glycans, predominantly expressed in human cells, over Neu5Gc-terminated

water-mediated hydrogen bonds with Tyr33, Ser35, Lys59, Thr65, and Arg100 in PltB (Figure 6D). In addition, the first Neu5Ac sugar ring makes hydrophobic contacts with the aromatic rings of Tyr33 and Tyr34. The second Neu5Ac contacts PltB through direct and water-mediated hydrogen bonds with Ser35, Asp36, Lys59, Asn61, Ser63, Ala130, and Thr131 of the glycan-binding domain (Figure S4). Consistent with their importance in carbohydrate binding, mutations in Tyr33, Ser35, and Lys59 drastically disrupted typhoid toxin activity (Figures 6E and S5). Some toxin B subunits undergo conformational changes upon binding their glycan receptors (Sixma et al., 1992). To investigate this possibility, we solved the atomic structure of the apo form of PltB at 2.08 Å resolution. Comparison of the atomic structures of the receptor-bound and apo forms of PltB indicates that binding to glycan receptors does not result in marked conformational changes in PltB (Figure S6).

Typhoid toxin's PltB shares its oligosaccharide-binding fold with the B subunits of other AB₅ toxins such as subtilase cytotoxin's SubB (Byres et al., 2008; Song et al., 2013). However, SubB exhibits the opposite specificity, strongly favoring binding to Neu5Gc-terminated glycans (Byres et al., 2008). Therefore, comparison of the atomic structures of PltB and SubB bound to their glycan receptors afforded us an opportunity to obtain insight into the structural bases for the binding specificity. The arrangement of the main chain of Neu5Ac relative to the binding pocket of PltB is very similar to that of Neu5Gc bound to SubB (Byres et al., 2008), and many of the critical interactions between the glycans and specific residues of PltB and SubB are conserved (Figure 6F). However, a residue equivalent to Tyr78 in SubB, which forms a critical hydrogen bond with the extra hydroxyl group in Neu5Gc is missing from PltB (Figure 6F). Instead, at this position PltB has the nonpolar residue Val103 and thus is unable to interact with Neu5Gc. These findings provide a structural explanation for typhoid toxin's inability to bind Neu5Gc-terminated glycans

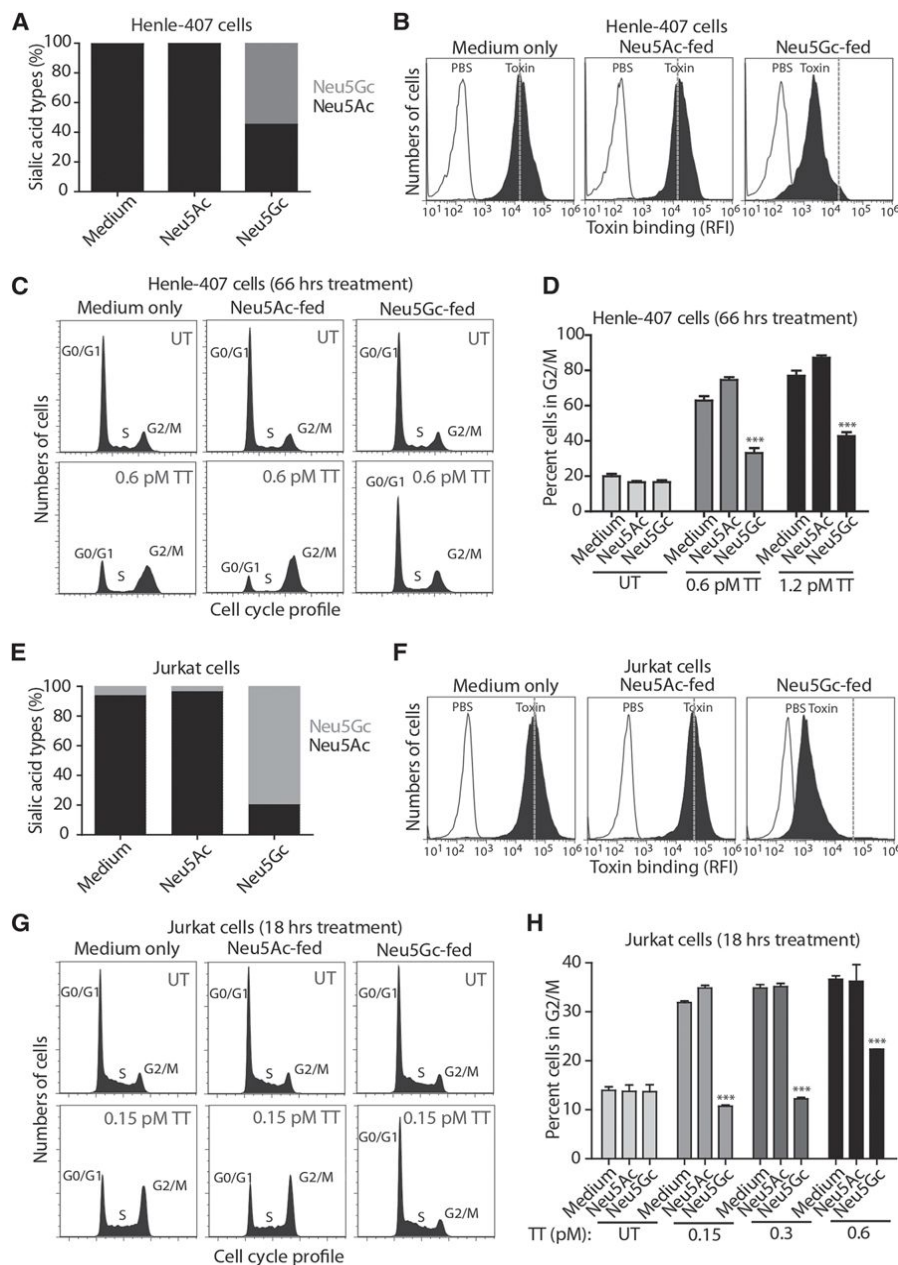


Figure 3. Typhoid Toxin Binds to and Is Cytotoxic toward Cells Displaying Neu5Ac-, But Not to Those Displaying Neu5Gc-, Terminated Glycans

Human intestinal epithelial Henle-407 cells (A–D) and human T lymphocyte Jurkat cells (E–H) were left untreated (medium only) or fed Neu5Ac or Neu5Gc in a culture medium for 4 days. Cells were then analyzed by HPLC to examine their relative sialic acid composition (A and E), or used in typhoid toxin binding (B and F) and toxicity assays by examining the cell cycle profile of toxin-treated cells (C, D, G, and H). Data in (D) and (H) are the mean \pm SEM; ***p < 0.0001, compared to the percent of control (medium-treated) cells in G2/M in the same group. See also Figure S2.

lase B subunit SubB bound to Neu5Gc with the structure of PltB bound to Neu5Ac revealed that the arrangement of the main chain of the two glycans relative to their binding pockets is very similar and many of the critical interactions between the glycans and specific residues of PltB and SubB are highly conserved. However, PltB lacks a residue equivalent to Tyr78 in SubB, which forms a critical hydrogen bond with the extra hydroxyl group in Neu5Gc. Because in all likelihood PltB evolved from a Neu5Gc-binding ancestor, this finding suggests that only subtle changes in the glycan-binding site would have been necessary to drastically change typhoid toxin's binding specificity and host range. However, mutagenesis analysis of PltB suggests a more complex picture since changing Val103 to Tyr in PltB (equivalent to Tyr78 in SubB) resulted in a loss of function rather than in a change in binding specificity (Figure S7). Additional structures of PltB bound to different glycans will be required to fully understand the evolution of typhoid toxin's exquisite binding specificity.

glycans, which are predominantly expressed by most other mammals. Therefore, the exquisite binding selectivity of typhoid toxin for glycans predominantly expressed in human cells provides an explanation for the inability of *S. Typhi* to cause typhoid fever in some nonpermissive species like chimpanzees, which allow significant bacterial replication.

The extreme specificity for human glycans exhibited by typhoid toxin is striking and unprecedented among bacterial toxins. The bacterial toxin subtilase, expressed by some strains of *E. coli*, exhibits the opposite specificity and its B subunit strongly favors binding to Neu5Gc-terminated glycans (Byres et al., 2008), which is consistent with the broad host specificity of this pathogen. Comparison of the crystal structures of subtilase

Nevertheless, from the host perspective, it is remarkable that a single oxygen atom could have such a dramatic impact on pathogenicity.

We have shown here that a toxin produced by a human-specific pathogen has evolved to selectively bind human sialoglycans. This is a remarkable example of virulence factor adaptation to a specific host that provides major insight into the process of host adaptation of the human pathogen *S. Typhi*. Given typhoid toxin's central role in the development of typhoid fever, these observations provide the bases for novel therapeutic strategies and may help the development of an animal model for the study of typhoid fever and the pathogenesis of typhoid toxin.

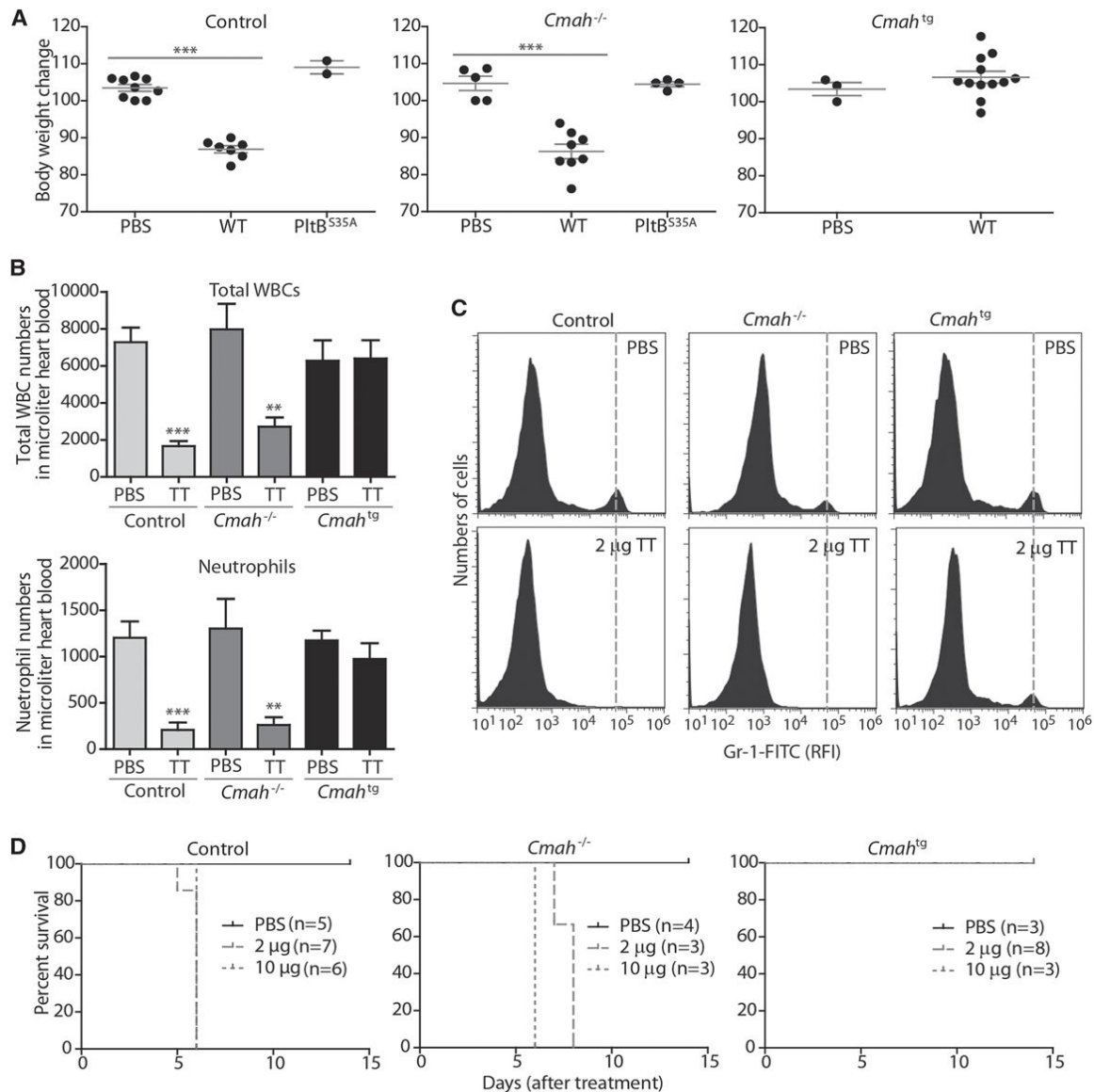


Figure 4. Mice Engineered to Constitutively Express CMAH, Resulting in Elevated Levels of Neu5Gc in All Tissues, Are Resistant to Typhoid Toxin

Purified preparations of wild-type typhoid toxin or a binding-defective mutant (PltB^{S35A}) were systemically administered into mice defective in (*Cmah*^{-/-}) or constitutively expressing CMAH (*Cmah*^{tg}) or control (C57BL/6) mice.

(A–C) Four days after treatment their total weight (A) and the total number of white cells (WBC) (B, top panel) or neutrophils (B, bottom) were measured as indicated in Experimental Procedures. Black circles represent the percentage of the weight of an animal relative to its weight immediately before treatment (A). Circulating white blood cells were counted in a hematology analyzer (B). Alternatively, peripheral blood cells from animals that had received the indicated treatments were stained with an antibody directed to the neutrophil cell marker Gr1 and the number of stained cells was determined by flow cytometry (C). The histograms shown are from ungated samples. Similar results were obtained in several independent repetitions of the experiment. RFI, relative fluorescence intensity. TT, typhoid toxin. WT, wild-type. Data in (B) are the mean \pm SEM; ***p < 0.0001, **p < 0.002 (relative to the buffer control in the same group).

(D) Survival of mice after administration of different amounts of typhoid toxin. PBS, phosphate buffered saline. The difference in the survival curves of PBS versus toxin-treated (all concentrations) control and *Cmah*^{-/-} animals was statistically significant (p < 0.001; log-rank Mantel-Cox test). The difference between the survival curves in control and *Cmah*^{-/-} mice after administration of 2 µg of toxin was statistically significant (p < 0.001). However, after administration of 10 µg of toxin the difference between the survival curves in control and *Cmah*^{-/-} mice was not statistically significant (p < 0.6).

See also Figure S3.

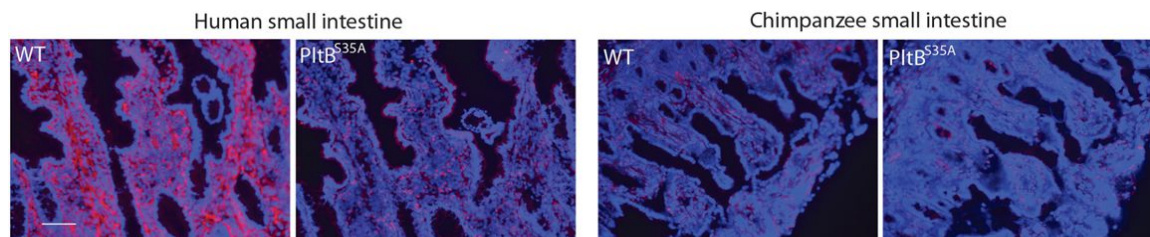


Figure 5. Typhoid Toxin Does Not Bind to Chimpanzee Tissues

Frozen sections of small intestine from humans or chimpanzees were stained with fluorescently labeled typhoid toxin or its binding-defective PltB^{S35A} mutant (red) and counterstained with Hoechst (blue). Scale bar represents 100 μ m.

EXPERIMENTAL PROCEDURES

Typhoid Toxin Expression and Purification

Expression and purification of typhoid toxin was carried out as previously described (Song et al., 2013). Plasmids expressing PltB point mutations were constructed using standard recombinant DNA techniques.

Crystallization

Expression and purification of C-terminal 6 \times His-tagged PltB used for crystallization have been described previously (Song et al., 2013). Initial sparse matrix crystallization trials of full-length PltB protein preparations (5.5 mg ml⁻¹) were carried out at the Yale University School of Medicine Structural Biology Core facility. After crystal optimization trials, full-length PltB crystals appeared in 2 to 3 days and matured in \sim 1 week at room temperature using the hanging-drop vapor-diffusion method in a mix of 1 μ l of protein with 1 μ l of reservoir solution consisting of 26% (w/v) PEG1500 and 0.1 M sodium acetate, pH5.0. Native PltB crystals were soaked by the addition of GD2 in different concentrations (from 1 mM to 50 mM). In most cases, the crystals broke upon addition of the sugar, even at sugar concentration of 1 mM. In very few cases, crystal debris large enough to be mounted on the X-ray source were obtained, which diffracted to \sim 3 Å.

X-Ray Data Collection and Structure Determination

All data were collected at a wavelength of 1.5418 Å on a Rigaku Homelab system at the Yale University Chemical and Biophysical Instrumentation Center (CBIC) (<http://cbic.yale.edu>). Data were integrated and scaled using the HKL-2000 package (Otwinowski and Minor, 1997). Further processing was performed with programs from the CCP4 suite (Collaborative Computational Project, Number 4, 1994). The apo and GD2 bound PltB structures were both determined by molecular replacement using PHASER (McCoy et al., 2007) with the atomic coordinates of chain A of typhoid toxin (Song et al., 2013) (Protein Data Bank [PDB] ID 4K6L) as the initial search model. To complete the model, manual building was carried out in COOT (Emsley and Cowtan, 2004). Figures were prepared using PyMol (DeLano, 2002). The structure refinement was done by PHENIX (Adams et al., 2010). The data collection and refinement statistics are summarized in Table S2. Coordinates for the atomic structures have been deposited in the RCSB Protein Data Bank under PDB numbers 4RHR and 4RHS.

Alexa 555 Typhoid Toxin Labeling

Purified wild-type and PltB^{S35A} mutant typhoid toxins were fluorescently labeled with Alexa-555 (Invitrogen) according to the vendor's recommendation. Purified typhoid toxin preparations (1 μ g/ml in 500 μ l of 100 mM bicarbonate buffer) were incubated for 1 hr at room temperature (RT) with reactive dye and applied to a size exclusion chromatography column to separate dye-protein conjugates from free dye. Degree of labeling was determined by measuring the absorbance of the conjugate solution at 280 and 555 nm. Efficiency of labeling was equivalent for both wild-type and PltB^{S35A} toxin preparations (4:1 dye/holotoxin ratio for both preparations). The typhoid holotoxin's predicted extinction coefficient is 191,400 M⁻¹ cm⁻¹.

Glycan Microarray Analysis

Glycan microarrays were fabricated using epoxide-derivatized glass slides as previously described (Padler-Karavani et al., 2014). Printed glycan microarray slides were blocked by ethanolamine, washed and dried, and then fitted in a multiwell microarray hybridization cassette (ArrayIt) to divide into subarrays. The subarrays were blocked with Ovalbumin (1% w/v) in PBS (pH 7.4) for 1 hr at RT in a humid chamber with gentle shaking. Subsequently, the blocking solution was discarded, and diluted wild-type or mutant typhoid toxin samples (Alexa Fluor 555-labeled) were added to each subarray. After incubating the toxins for 2 hr at RT with gentle shaking, the slides were extensively washed to remove nonspecifically bound proteins. The developed glycan microarray slides were then dried and subjected to scanning by a Genepix 4000B microarray scanner (Molecular Devices) immediately. Data analysis was done using the Genepix Pro 7.0 analysis software (Molecular Devices).

Typhoid Toxin-Sialoglycan Binding Affinity Measured by Microscale Thermophoresis

The binding affinity of typhoid toxin to different glycans was measured using microscale thermophoresis as previously described (Wienken et al., 2010). Briefly, NT-647 fluorescently labeled typhoid toxin was incubated with a wide range of concentrations (12 μ M to 400 mM over 15 2-fold serial dilutions) of Neu5Ac α 2-3Gal β 1-4Glc, at room temperature for 30 min in the dark. After equilibrium, the mixtures were loaded into 16 hydrophilic glass capillaries and the microscale thermophoresis analysis was performed using the Monolith NT.115 (Nano Temper). Data were analyzed and binding affinities were determined by the Nano Temper analysis software package. The sialoglycan underlying structure, Gal β 1-4Glc (lactose), was used as a control under the same conditions.

Mammalian Cell Culture Conditions

Henle-407 human intestinal epithelial and Jurkat human T lymphocyte cells were cultured in DMEM high glucose + 10% FBS and RPMI1640 + 10% FBS + 1 mM sodium pyruvate + 10 mM HEPES, respectively. All mammalian cells were kept at 37°C in a cell culture incubator with 5% CO₂. Metabolic incorporation of Neu5Ac or Neu5Gc was carried out as previously described (Tangvoranuntakul et al., 2003). Briefly, cells were cultured in a standard medium supplemented with 10 mM Neu5Ac or 10 mM Neu5Gc (Inalco) as follows. Henle-407 (1.5×10^4) or Jurkat (1×10^5) cells were seeded into 12-well culture plates in 1 ml media with or without 10 mM sialic acid. A stock solution of 50 mM sialic acid was freshly prepared in a DMEM medium whose final pH was adjusted to neutral with NaOH. During the feeding period, the cells were continuously monitored and maintained below 80% confluence. After 3 days of growth, cells were split into 12-well plates at a cell density of 1.5×10^4 (Henle-407) or 1×10^5 (Jurkat) cells per well. The next day the cells were used for different assays as described below.

High-Performance Liquid Chromatography

High-performance liquid chromatography (HPLC) analysis of sialic acids was carried out as previously described (Tangvoranuntakul et al., 2003). Briefly, sialic acids were released from glycans by hydrolysis with acetic acid, filtered

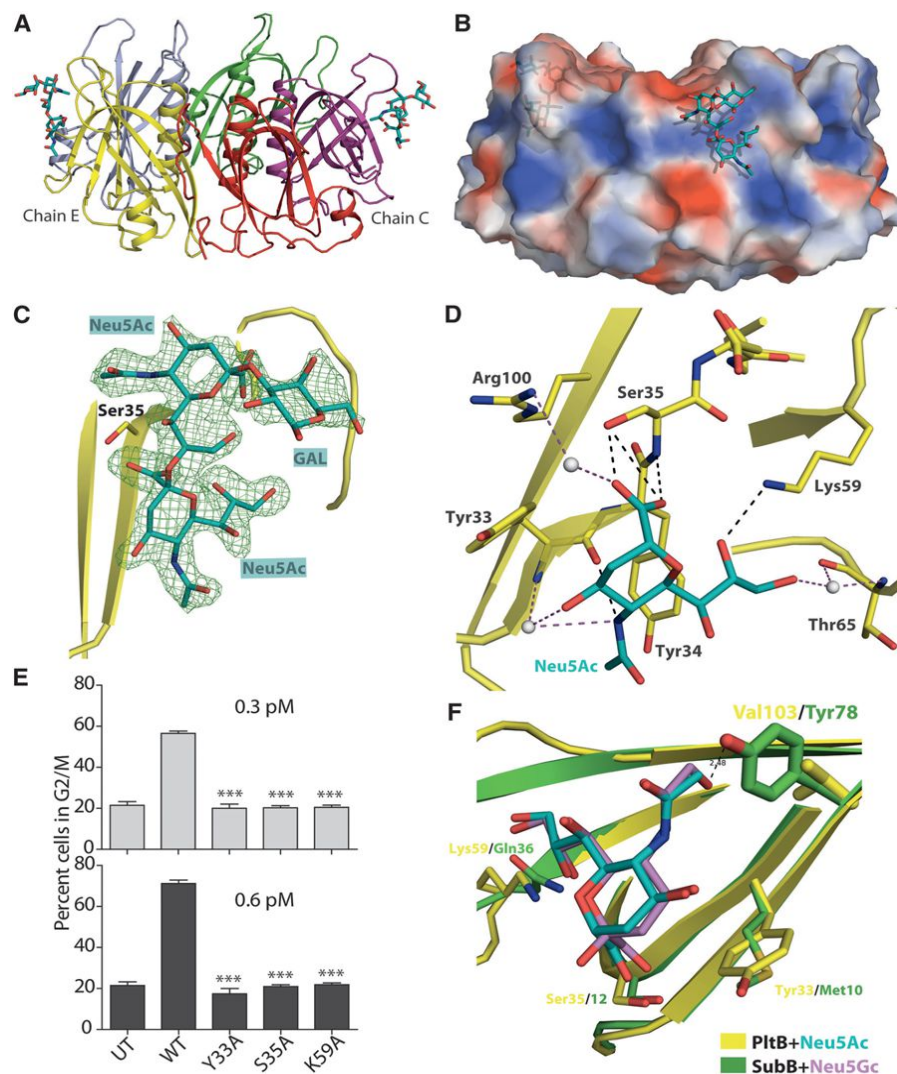


Figure 6. Crystal Structure of Typhoid Toxin B Subunit PltB Bound to Its Sialic Acid Ligand

(A) The atomic structure of the PltB pentamer in complex with the GalNAcβ1-4(Neu5Acα2-8Neu5Acα2-3)Galβ1-4Glc oligosaccharide is shown as a ribbon cartoon with each protomer depicted in a different color. In the PltB pentamer, only partial oligosaccharide density (Neu5Ac-Neu5Ac-Gal) is seen in Chain C (purple) and E (yellow). Cyan sticks represent the sugar carbon atoms, blue sticks represent nitrogen atoms, and red sticks represent oxygen atoms.

(B) Surface charge distribution of the PltB pentamer structure and sugar-binding pockets.

(C) Close-up views of Neu5Ac-Neu5Ac-Gal and its composite annealed omit difference density map. PltB chain E and its key residue Ser35 are shown in yellow. Green mesh represents the sugar difference density map contoured at 2.5σ.

(D) Interactions between PltB and Neu5Ac. Chain E of PltB is shown as a yellow colored ribbon cartoon, the amino acids interacting with the sugar are shown as sticks, and the direct interactions are shown in black dash. Water is shown as gray balls and water-mediated interactions are shown as purple dashes.

(E) Structure/function analysis of the PltB glycan-binding site. Typhoid holotoxin toxin preparations containing the indicated PltB mutants were tested for their ability to intoxicate cultured Henle-407 cells. Toxicity was evaluated by determining the percentage of cells arrested at the G2/M phase of the cell cycle, which is a measure of typhoid toxin's CdtB activity. Data are the mean ± SEM; ***p < 0.0001, compared to the percent cells treated with wild-type toxin that are in G2/M.

(F) Comparison of the sugar binding sites of PltB and SubB bound to Neu5Ac and Neu5Gc, respectively. Critical residues that differ between SubB (Tyr78) and PltB (Val103) are highlighted as sticks. Other interacting amino acids and sugars are shown in lines. PltB is shown in yellow, Neu5Ac in Cyan, SubB in Green, and Neu5Gc in light purple. See also Figures S4, S5, S6, and S7 and Table S2.

through a 10K Microcon filter unit (Millipore) and then derivatized by 1,2-diamino-4,5-methylene dioxybenzene (DMB) at 50°C for 2.5 hr in the dark. Resulting samples were analyzed by HPLC using a C18-column.

Typhoid Toxin Binding Assay

Cultured cells grown under different conditions (see above) were harvested, washed with Hank's balanced salt solution (HBSS), and resuspended in 100 μl HBSS containing 0.3 (for cultured Henle-407 or Jurkat cells) or 0.5 μg (for human and chimpanzee primary cells) of Alexa 555-labeled wild-type or mutant toxin preparations. Cells were incubated in the presence of the labeled toxin preparations for 15 min on ice and immediately analyzed by flow cytometry. The binding profiles were analyzed using Flowjo (Treestar).

Mammalian Cell Intoxication Assay

Cell-cycle arrest after typhoid toxin intoxication was examined by flow cytometry as previously described (Spanò et al., 2008). Briefly, after treatment with 6xHis-tagged typhoid toxin for 66 hr for Henle-407 or 18 hr for Jurkat cells, cells were trypsinized, collected, washed, and fixed for 2 hr in ~70% ethanol/PBS at -20°C. Fixed cells were washed with PBS and resuspended in 500 μl of PBS containing 50 μg/ml propidium iodide, 0.1 mg/ml RNase A,

and 0.05% Triton X-100. After incubation for 40 min at 37°C, cells were washed with PBS, resuspended in 500 μl PBS, filtered, and analyzed by a flow cytometry. The DNA content of treated cells was determined using Flowjo program (Treestar).

Mouse Intoxication Experiments

All mouse experiments were conducted according to protocols approved by Yale University's Institutional Animal Care and Use Committee. Age- and sex-matched 5- to 7-week-old C57BL/6 (wild-type), *Cmah*^{-/-} (Jackson Laboratory), or *Cmah*¹⁹ mice were intravenously injected with 100 μl solutions containing PBS alone, or either 2 μg or 10 μg of the indicated purified toxin preparations. Changes in behavior, weight, and survival of the toxin-injected mice were closely monitored and to minimize bias, blind end-point assessment was applied to all the experiments.

Blood Counting

Blood samples were collected by heart puncture 4.5 days after toxin treatment in Microtainer tubes coated with EDTA as an anticoagulant (BD), kept at room temperature, and analyzed within 2 hr after blood collection using a Hemavet 950FS hematology analyzer (Drew Scientific). Blood counts were analyzed by GraphPad Prism (GraphPad Software).

Peripheral Blood Cell Preparation, Immunostaining, and Flow Cytometry Analysis

Peripheral blood samples of typhoid toxin-treated and control mice were collected into tubes coated with EDTA, incubated with 1 ml ACK buffer (BioWhittaker), incubated for 5 min, washed with 2 ml PBS, and centrifuged to collect peripheral blood leukocytes (PBLs). After a repetition of the red blood cell removal step, PBLs were washed, and were immediately incubated for 30 min on ice with 100 μ l of anti-mouse Ly-6G (Gr-1) antibody conjugated with FITC (eBioscience, 11-5931-81). PBLs were then washed with 2 ml of FACS buffer (PBS, 0.16% BSA), resuspended in 100 μ l FACS fixation buffer (PBS, 1% paraformaldehyde, 1% FCS), and used for flow cytometric analyses on BD accuri C6 (BD Biosciences). Peripheral blood samples from humans and chimpanzees were collected into EDTA tubes. Erythrocytes were separated from peripheral blood monocyctic cells (PBMC) by Ficoll-Paque Plus. Erythrocytes in the PBMC layer were lysed by ACK buffer, and monocytes were removed by anti-CD14 beads (MACS Miltenyi Biotec).

Typhoid Toxin Binding to Human, Chimpanzee, and Mouse Tissues

Cryosections of frozen tissue samples from human, chimpanzee and the different mouse strains, were overlaid with AF555-labeled wild-type typhoid toxin or the PltB^{S35A} mutant. After incubation in a covered humid chamber for 1 hr at room temperature, the slides were washed to remove nonbound toxins and the sections were fixed using 10% neutral buffered formalin. The nuclei were counterstained using Hoechst and the slides were washed and mounted in aqueous mounting media (VectaMount). Digital photomicrographs were taken using a Keyence BZ9000 fluorescence microscope (BIOREVO, BZ-9000, Keyence).

Statistical Analysis

Two-tailed Student's *t* tests were performed in order to determine the statistical significance of experimental changes from control values. A *p* value < 0.05 was considered as statistically significant.

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2014.10.057>.

AUTHOR CONTRIBUTIONS

L.D. conducted or contributed to experiments shown in Figures 1, 2, 3A, 3E, 5, S1, and S3 and Tables 1 and S1. J.S. conducted or contributed to experiments shown in Figures 1, 2B, 2D, 3B–3D, 3F–3H, 4A–4D, 6E, S2, S5, and S7 and Tables 1 and S1. X.G. conducted experiments shown in Figures 6A–6D, 6F, S4, and S6 and Table S2. J.W. contributed to the analysis of the crystal structures. X.C. and H.Y. contributed to experiments shown in Figure 1 and Tables 1 and S1. N.V. planned, supervised, and interpret experiments shown in Figures 5 and S3. Y.N.-M. contributed to experiments shown in Figure 4D. J.E.G. and A.V. contribute to the design, interpretation, and supervision of this study. J.E.G. wrote the paper with input from A.V. and comments from all authors.

ACKNOWLEDGMENTS

We thank Jonathon Okerblom for help with sialic acid analysis, Ana Lazic for assistance in the microscale thermophoresis studies, and members of the Galán laboratory for critical review of the manuscript. This work was supported by the National Cancer Institute (NCI) grant CA38701 to A.V. and National Institute of Allergy and Infectious Diseases (NIAID) grant AI079022 to J.E.G.

Received: August 8, 2014

Revised: September 9, 2014

Accepted: October 28, 2014

Published: December 4, 2014

REFERENCES

- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221.
- Beddoe, T., Paton, A.W., Le Nours, J., Rossjohn, J., and Paton, J.C. (2010). Structure, biological functions and applications of the AB5 toxins. *Trends Biochem. Sci.* **35**, 411–418.
- Byres, E., Paton, A.W., Paton, J.C., Löfling, J.C., Smith, D.F., Wilce, M.C., Talbot, U.M., Chong, D.C., Yu, H., Huang, S., et al. (2008). Incorporation of a non-human glycan mediates human susceptibility to a bacterial toxin. *Nature* **456**, 648–652.
- Chou, H.H., Hayakawa, T., Diaz, S., Krings, M., Indriati, E., Leakey, M., Paabo, S., Satta, Y., Takahata, N., and Varki, A. (2002). Inactivation of CMP-N-acetylneuraminic acid hydroxylase occurred prior to brain expansion during human evolution. *Proc. Natl. Acad. Sci. USA* **99**, 11736–11741.
- Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763.
- Crump, J.A., and Mintz, E.D. (2010). Global trends in typhoid and paratyphoid fever. *Clin. Infect. Dis.* **50**, 241–246.
- DeLano, W.L. (2002). The PyMOL Molecular Graphics System. <http://www.pymol.org>.
- Edsall, G., Gaines, S., Landy, M., Tigertt, W.D., Sprinz, H., Trapani, R.J., Mandel, A.D., and Benenson, A.S. (1960). Studies on infection and immunity in experimental typhoid fever. I. Typhoid fever in chimpanzees orally infected with *Salmonella typhosa*. *J. Exp. Med.* **112**, 143–166.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132.
- Galán, J.E. (2001). *Salmonella* interactions with host cells: type III secretion at work. *Annu. Rev. Cell Dev. Biol.* **17**, 53–86.
- Haghjoo, E., and Galán, J.E. (2004). *Salmonella typhi* encodes a functional cytolethal distending toxin that is delivered into host cells by a bacterial-internalization pathway. *Proc. Natl. Acad. Sci. USA* **101**, 4614–4619.
- Hedlund, M., Tangvoranuntakul, P., Takematsu, H., Long, J.M., Housley, G.D., Kozutsumi, Y., Suzuki, A., Wynshaw-Boris, A., Ryan, A.F., Gallo, R.L., et al. (2007). N-glycolylneuraminic acid deficiency in mice: implications for human biology and evolution. *Mol. Cell. Biol.* **27**, 4340–4346.
- Ibarra, J.A., and Steele-Mortimer, O. (2009). *Salmonella*—the ultimate insider. *Salmonella* virulence factors that modulate intracellular survival. *Cell. Microbiol.* **11**, 1579–1586.
- Jacobsen, A., Hendriksen, R.S., Aarestrup, F.M., Ussery, D.W., and Friis, C. (2011). The *Salmonella enterica* pan-genome. *Microb. Ecol.* **62**, 487–504.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674.
- Metchnikoff, E., and Besredka, A. (1911). Recherches sur la fièvre typhoïde expérimentale. *Ann. Inst. Pasteur (Paris)* **25**, 865–874.
- Ohl, M.E., and Miller, S.I. (2001). *Salmonella*: a model for bacterial pathogenesis. *Annu. Rev. Med.* **52**, 259–274.
- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
- Padler-Karavani, V., Hurtado-Ziola, N., Chang, Y.C., Sonnenburg, J.L., Ronaghy, A., Yu, H., Verhagen, A., Nizet, V., Chen, X., Varki, N., et al. (2014). Rapid evolution of binding specificities and expression patterns of inhibitory CD33-related Siglecs in primates. *FASEB J.* **28**, 1280–1293.
- Parkhill, J., Dougan, G., James, K.D., Thomson, N.R., Pickard, D., Wain, J., Churcher, C., Mungall, K.L., Bentley, S.D., Holden, M.T., et al. (2001). Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. *Nature* **413**, 848–852.

- Parry, C.M., and Threlfall, E.J. (2008). Antimicrobial resistance in typhoidal and nontyphoidal salmonellae. *Curr. Opin. Infect. Dis.* 21, 531–538.
- Parry, C.M., Hien, T.T., Dougan, G., White, N.J., and Farrar, J.J. (2002). Typhoid fever. *N. Engl. J. Med.* 347, 1770–1782.
- Raffatellu, M., Wilson, R.P., Winter, S.E., and Bäumler, A.J. (2008). Clinical pathogenesis of typhoid fever. *J. Infect. Dev. Ctries.* 2, 260–266.
- Sabbagh, S.C., Forest, C.G., Lepage, C., Leclerc, J.M., and Daigle, F. (2010). So similar, yet so different: uncovering distinctive features in the genomes of *Salmonella enterica* serovars Typhimurium and Typhi. *FEMS Microbiol. Lett.* 305, 1–13.
- Schwan, W.R., Huang, X.Z., Hu, L., and Kopecko, D.J. (2000). Differential bacterial survival, replication, and apoptosis-inducing ability of *Salmonella* serovars within human and murine macrophages. *Infect. Immun.* 68, 1005–1013.
- Sixma, T.K., Pronk, S.E., Kalk, K.H., van Zanten, B.A., Berghuis, A.M., and Hol, W.G. (1992). Lactose binding to heat-labile enterotoxin revealed by X-ray crystallography. *Nature* 355, 561–564.
- Song, J., Gao, X., and Galán, J.E. (2013). Structure and function of the *Salmonella* Typhi chimeric A(2)B(5) typhoid toxin. *Nature* 499, 350–354.
- Spanò, S., and Galán, J.E. (2012). A Rab32-dependent pathway contributes to *Salmonella* typhi host restriction. *Science* 338, 960–963.
- Spanò, S., Ugalde, J.E., and Galán, J.E. (2008). Delivery of a *Salmonella* Typhi exotoxin from a host intracellular compartment. *Cell Host Microbe* 3, 30–38.
- Spanò, S., Liu, X., and Galán, J.E. (2011). Proteolytic targeting of Rab29 by an effector protein distinguishes the intracellular compartments of human-adapted and broad-host *Salmonella*. *Proc. Natl. Acad. Sci. USA* 108, 18418–18423.
- Srikanth, C.V., Mercado-Lubo, R., Hallstrom, K., and McCormick, B.A. (2011). *Salmonella* effector proteins and host-cell responses. *Cell. Mol. Life Sci.* 68, 3687–3697.
- Tangvoranuntakul, P., Gagneux, P., Diaz, S., Bardor, M., Varki, N., Varki, A., and Muchmore, E. (2003). Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *Proc. Natl. Acad. Sci. USA* 100, 12045–12050.
- Varki, N.M., Strobert, E., Dick, E.J., Jr., Benirschke, K., and Varki, A. (2011). Biomedical differences between human and nonhuman hominids: potential roles for uniquely human aspects of sialic acid biology. *Annu. Rev. Pathol.* 6, 365–393.
- Vladoianu, I.R., Chang, H.R., and Pechère, J.C. (1990). Expression of host resistance to *Salmonella* typhi and *Salmonella* typhimurium: bacterial survival within macrophages of murine and human origin. *Microb. Pathog.* 8, 83–90.
- Voetsch, A.C., Van Gilder, T.J., Angulo, F.J., Farley, M.M., Shallow, S., Marcus, R., Cieslak, P.R., Deneen, V.C., and Tauxe, R.V.; Emerging Infections Program FoodNet Working Group (2004). FoodNet estimate of the burden of illness caused by nontyphoidal *Salmonella* infections in the United States. *Clin. Infect. Dis.* 38 (Suppl 3), S127–S134.
- Waterman, S.R., and Holden, D.W. (2003). Functions and effectors of the *Salmonella* pathogenicity island 2 type III secretion system. *Cell. Microbiol.* 5, 501–511.
- Wienken, C., Baaske, P., Rothbauer, U., Braun, D., and Duhr, S. (2010). Protein-binding assays in biological liquids using microscale thermophoresis. *Nat. Commun.* 1, 100.

Beta-Lactam Antibiotics Induce a Lethal Malfunctioning of the Bacterial Cell Wall Synthesis Machinery

Hongbaek Cho,¹ Tsuyoshi Uehara,^{1,2} and Thomas G. Bernhardt^{1,*}

¹Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA 02115, USA

²Present address: Novartis Institutes for Biomedical Research, Emeryville, CA 94608, USA

*Correspondence: thomas_bernhardt@hms.harvard.edu

<http://dx.doi.org/10.1016/j.cell.2014.11.017>

SUMMARY

Penicillin and related beta-lactams comprise one of our oldest and most widely used antibiotic therapies. These drugs have long been known to target enzymes called penicillin-binding proteins (PBPs) that build the bacterial cell wall. Investigating the downstream consequences of target inhibition and how they contribute to the lethal action of these important drugs, we demonstrate that beta-lactams do more than just inhibit the PBPs as is commonly believed. Rather, they induce a toxic malfunctioning of their target biosynthetic machinery involving a futile cycle of cell wall synthesis and degradation, thereby depleting cellular resources and bolstering their killing activity. Characterization of this mode of action additionally revealed a quality control function for enzymes that cleave bonds in the cell wall matrix. The results thus provide insight into the mechanism of cell wall assembly and suggest how best to interfere with the process for future antibiotic development.

INTRODUCTION

Penicillin and related beta-lactam drugs are one of our oldest and most widely used antibiotic classes. They have long been known to interfere with bacterial cell wall assembly as part of their mode of action (Park and Strominger, 1957). The cell wall is an essential polysaccharide structure that surrounds most bacterial cells and protects their cytoplasmic membrane from osmotic rupture. It is built from the polymer peptidoglycan (PG), which consists of glycan chains with attached peptides used to crosslink adjacent glycans to form a matrix structure (Figure 1A).

Beta-lactams disrupt PG biogenesis by inactivating enzymes called penicillin-binding proteins (PBPs) (Tipper and Strominger, 1965). Bacteria encode a variety of PBPs that participate in PG assembly (Sauvage et al., 2008). The high-molecular-weight PBPs are the major PG synthases. They are subdivided into class A (aPBPs) and class B (bPBPs) enzymes (Figure 1B). aPBPs are bifunctional and possess both glycosyltransferase (GT) activity for polymerizing the glycan strands and transpepti-

dase (TP) activity for crosslinking them. bPBPs, on the other hand, are only known to possess TP activity. The primary target of beta-lactams is the TP active site of the synthetic PBPs, which is covalently modified by the drug. In addition to the PG synthases, beta-lactams also inhibit the low-molecular-weight PBPs. These factors belong to a large and diverse family of enzymes that cleave bonds in the PG matrix. Such enzymes, often referred to as PG hydrolases, are typically nonessential, but have been found to play important roles in morphogenesis (Uehara and Bernhardt, 2011).

The lethal activity of beta-lactams is thought to stem principally from the loss of wall integrity accompanied by cell lysis (Park and Strominger, 1957). According to the most widely accepted model, cell wall damage following beta-lactam treatment results from a drug-induced imbalance between the activities of cell wall synthases and hydrolases (Schwarz et al., 1969; Tomasz and Waks, 1975; Tomasz et al., 1970). This view is supported by the observation that PG hydrolase inactivation can prevent or delay beta-lactam-induced cell lysis (Chung et al., 2009; Heidrich et al., 2002; Tomasz, 1979; Tomasz and Waks, 1975; Tomasz et al., 1970; Uehara et al., 2009). However, surprisingly little mechanistic insight underlies this general framework for drug action. It remains largely unclear which PG hydrolases disrupt the wall following drug treatment and whether these autolysins are “induced” to damage the wall or are simply carrying out their normal physiological function in the absence of TP activity. Clues suggesting a more complex mode of action for beta-lactams than simple TP inhibition have also been reported. Surprisingly, in *Streptococcus pneumoniae* mutants blocked for cell lysis, beta-lactam treatment still promoted cell death with kinetics similar to lysing cells (Moreillon et al., 1990). Additionally, in *Escherichia coli*, beta-lactams were unexpectedly shown to cause a depletion of lipid-linked cell wall precursors prior to cell lysis (Kohlrausch and Höltje, 1991).

The lysis-independent effects of beta-lactams indicate that much remains to be learned about the events following TP inhibition and how they contribute to the killing activity of this vital class of antibiotics. These events have been difficult to elucidate because the effect of drug treatment on cell growth and morphology varies depending on the organism studied and on the number of different PBPs inhibited by the beta-lactam derivative used (Spratt, 1975). We overcame this challenge by studying the activity of mecillinam, a beta-lactam specific for a single essential PBP in the model gram-negative bacterium *Escherichia coli*

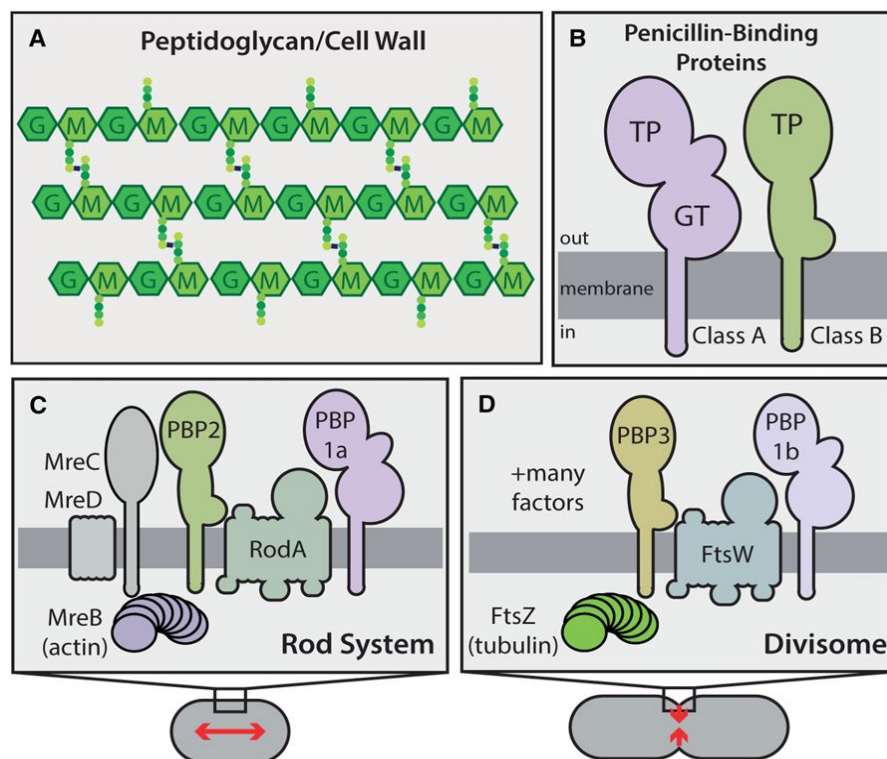


Figure 1. Peptidoglycan Structure and the Machines that Synthesize It

(A) The PG matrix consists of glycan chains with the repeating unit of N-acetylmuramic acid (MurNAc, M) and N-acetylglucosamine (GlcNAc, G). Attached to the MurNAc sugars are peptides (colored circles) used to form crosslinks between adjacent glycans.

(B) Domain structure of the PG synthases. Both classes of PBPs have a single transmembrane domain with a large catalytic domain in the periplasm.

(C and D) Schematics highlighting the components of the two main PG synthetic systems in rod-shaped bacteria. Both systems require a dedicated bPBP (PBP2 or PBP3) and a SEDS (shape/elongation/division/sporulation) family protein (RodA or FtsW) for proper function. See text for details.

(Spratt, 1975). Our analysis revealed that, beyond simply inhibiting the TP activity of PBPs, mecillinam and other beta-lactams stimulate a deleterious futile cycle of cell wall synthesis and degradation by their target machineries that contributes to their lethal activity. Additional genetic analysis identified the enzyme responsible for beta-lactam-stimulated degradation of nascent PG. Characterization of the *in vivo* activity of this factor suggests a novel “quality control” function for cell wall cleaving enzymes in PG biogenesis. Our findings thus provide new insight into the cell wall assembly process in addition to uncovering an important mechanism by which beta-lactam antibiotics induce cell death.

RESULTS

Rationale

Like many rod-shaped bacteria, *E. coli* grows using two different PG biogenesis systems (Typas et al., 2012) (Figures 1C and 1D). The actin-like MreB protein and its partners constitute the Rod system, which catalyzes the insertion of new PG material along the cell body to promote cell elongation (Typas et al., 2012) (Figure 1C). The tubulin-like FtsZ protein, on the other hand, organizes the divisome to synthesize PG for the new daughter cell poles (de Boer, 2010) (Figure 1D). Each of these machineries requires an essential bPBP for their activity: PBP2 for the Rod system and PBP3 for the divisome (Typas et al., 2012) (Figures 1B–D). Proper PG biogenesis by these systems in *E. coli* is also thought to require the aPBPs, PBP1a and PBP1b, to promote glycan chain polymerization and crosslinking (Typas et al., 2012) (Figures 1B–1D).

with the Rod⁺ FtsZ^{up} cells growing and dividing as small spheres (Bendezú and de Boer, 2008; Kruse et al., 2005; Vinella et al., 1992). However, the mechanism by which increased FtsZ concentration suppresses the lethality of Rod system inactivation is not clear. In addition to genetic inactivation, Rod system function can also be blocked with the small molecules A22 (Gitai et al., 2005) or mecillinam (Spratt, 1975) (Figure 2A). A22 inactivates MreB (Gitai et al., 2005), and mecillinam is a beta-lactam that is specific for the TP active site of PBP2 (Spratt, 1975). Because PBP2 can be rendered nonessential by the overproduction of FtsZ (Vinella et al., 1993) (Figure 2B), we reasoned that studies of mecillinam presented a unique opportunity to study the consequences of PBP inhibition by beta-lactams under conditions where the target is not required for growth.

Mecillinam Induces a Lethal Malfunctioning of the Rod System

The lethal effects of inactivating MreB with A22 are suppressed by the low-copy vector pTB63, from which the *ftsQAZ* operon is expressed via its native promoters (a condition we will henceforth designate as FtsZ^{up}) (Bendezú and de Boer, 2008) (Figure 2C). Surprisingly, however, FtsZ^{up} cells remained sensitive to mecillinam (Figure 2C). Thus, blocking PBP2 function with mecillinam is lethal even under conditions where the protein itself is not essential (Figure 2B). This observation suggested that, beyond inactivation of the TP activity of PBP2, mecillinam is also conferring a dominant-negative function to the inactivated protein. We reasoned that this toxic activity of inhibited PBP2 may stem from the malfunctioning of the Rod system. To test this hypothesis, we treated cells with both mecillinam and A22

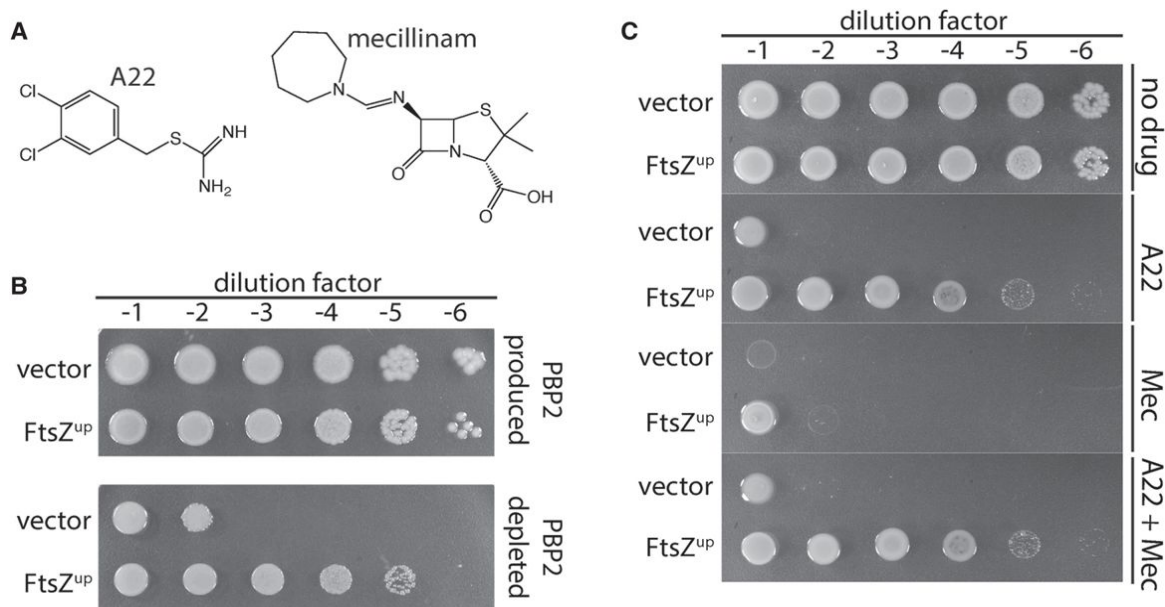


Figure 2. Mecillinam Induces a Lethal Malfunctioning of the Rod System

(A) Shown are the chemical structures of A22 and mecillinam.

(B) Conditional essentiality of PBP2. Cultures of strain HC439/pHC817 [$\Delta pbpA$ $P_{lac}::pbpA$] harboring either a vector control (pSC101, vector) or a derivative (pTB63, FtsZ^{up}) were serially diluted and spotted on LB agar with either 1 mM IPTG (PBP2 replete) or agar without inducer (PBP2 depletion). Plates were incubated overnight at 30°C and photographed. **C.** Cultures of strain MG1655 [WT] harboring either pSC101 or pTB63 were serially diluted and spotted on LB agar as in part B. Agar was supplemented with A22 (10 μg/ml), mecillinam (1 μg/ml), or both drugs as indicated.

to simultaneously inactivate the Rod system and the TP domain of PBP2. Indeed, inactivation of MreB with A22 in FtsZ^{up} cells completely suppressed the toxicity of mecillinam (Figure 2C). Genetic inactivation of the Rod system was also found to restore growth to FtsZ^{up} cells treated with mecillinam (Figure S1 available online). The contribution of Rod system malfunction to the lethal action of mecillinam was also apparent for cells lacking pTB63. Addition of mecillinam to a culture of wild-type cells led to a drop in culture OD₆₀₀ and a loss of viability that was reduced in severity by cotreatment with A22 (Figures S2A and S2B). The production of a PBP2 variant with a defective TP active site [PBP2(S330A)] also killed FtsZ^{up} cells (Figure S2C), and this toxicity could also be suppressed by A22 treatment (Figure S2C). Thus, the toxicity induced by mecillinam is specific for PBP2 inhibition, and not the result of an off-target activity. We conclude that blocking the TP activity of PBP2 causes the Rod system to malfunction in a manner that contributes to the lethal activity of mecillinam.

A Futile Cycle of PG Synthesis and Turnover Is Responsible for Mecillinam Toxicity

Results from Uehara and Park (2008) previously found that mecillinam treatment induced the degradation of newly synthesized PG while A22 blocks synthesis. We suspected that the observed turnover of nascent PG was key to the lethal mode of action of mecillinam. We therefore monitored PG synthesis and turnover using similar procedures to those of Uehara and Park (2008). Cells were metabolically labeled with [³H]-meso-diaminopimelic acid (mDAP), an amino acid unique to the peptide moiety of PG

(Figure 3A), following which, PG synthesis can be readily measured by monitoring the production of the major cytoplasmic precursor UDP-MurNAc-pentapeptide (UDP-MurNAc-pep₅) and the incorporation of label into the PG matrix. [³H]-mDAP labeling can also be used to quantify PG turnover. The major degradation products of PG are disaccharide fragments liberated by lytic transglycosylases (LTs) (Fisher and Mobashery, 2014) (Figure 3A). Rather than hydrolyzing the glycosidic bond, these enzymes promote the formation of an anhydro linkage between the C1 and C6 positions of MurNAc to generate GlcNAc-anhydro-MurNAc-peptides (GlcNAc-anhMurNAc-peptide). These degradation products are normally imported and processed by a number of enzymes in the cytoplasm for recycling (Fisher and Mobashery, 2014) (Figure 3A). We therefore used a strain lacking the amidase AmpD to trap PG turnover products and allow the quantification of PG degradation without complications of complete recycling (Uehara and Park, 2008) (Figure 3A).

To specifically monitor the effect of drug treatment on PG synthesis and degradation by the Rod system, divisome activity was blocked during the radiolabeling by expression of the FtsZ antagonist Sula prior to the addition of antibiotic(s). After 5 min of growth in the presence of drugs, [³H]-mDAP was added to the medium and growth was continued for 10 min (1/10th of a mass doubling). Cells were then harvested and extracted with hot (90°C) water. Soluble compounds (PG precursors and degradation products) were separated and quantified by HPLC and radiodetection. Material incorporated into the PG matrix was quantified as radioactive material released from the cell pellet by lysozyme treatment.

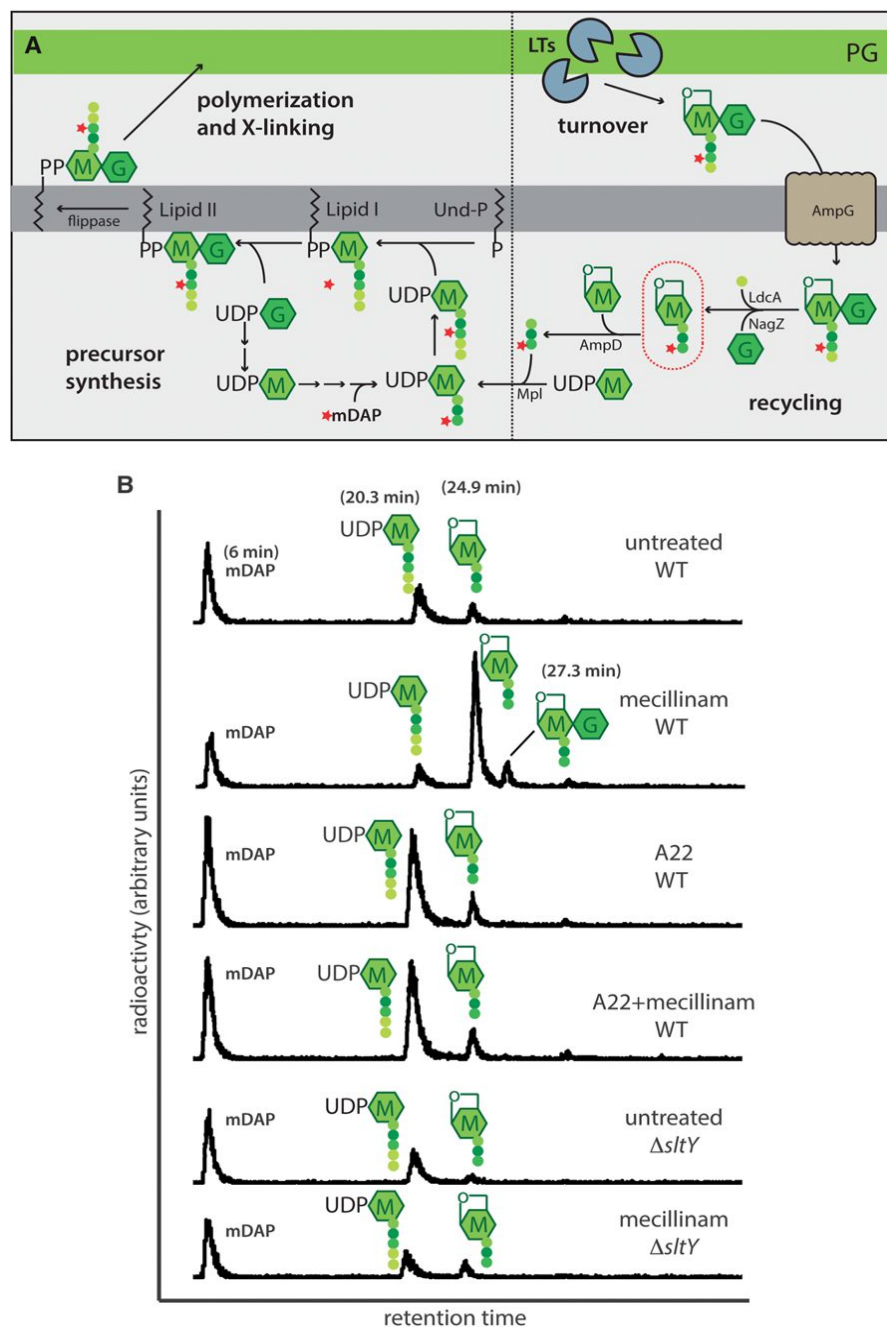


Figure 3. Measurement of PG Synthesis and Turnover following Mecillinam Treatment

(A) Schematic summarizes the PG synthesis and recycling pathways. Sugars and peptides are represented as in Figure 1A. The composition of the pentapeptide of the cytoplasmic precursors is L-Ala-gamma-D-Glu-meso-diaminopimelic acid (mDAP)-D-Ala-D-Ala. UDP-sugars are first made in the cytoplasm before being transferred to the lipid carrier undecaprenol-phosphate (Und-P). The final precursor lipid II is the substrate used by the PBPs to form PG. As a result of the crosslinking mechanism and the activity of carboxypeptidase enzymes that remove the terminal D-Ala, the PG layer consists primarily of glycans with tetrapeptides. The turnover product circled in red will accumulate in a $\Delta ampD$ strain if nascent PG is degraded during a pulse labeling experiment. See text for details.

(B) Measurement of PG synthesis and turnover by the Rod system. Cells of TU278 [$\Delta lysA \Delta ampD$] or its Δslt derivative (HC419) as indicated producing SuIA to block cell division were pulse labeled with [3H]-mDAP following treatment with the indicated drug(s) (10 μ g/ml each when added). Soluble metabolites were separated by HPLC and detected using an in-line radiodetector. The resulting chromatograms are shown. Peaks are labeled with schematics of the corresponding compounds. The identity of each peak is based on the results shown in Figure S3. See text for details.

block in the synthesis pathway. Only a modest increase in anh MurNAc-peptide accumulation was observed following A22 treatment, indicating that the drop in label incorporation into the PG matrix resulted primarily from the inhibition of PG synthesis. Similarly, in cells treated with mecillinam, incorporation of label into PG was also greatly reduced (Figures 3B and 4). However, as expected from the results of Uehara and Park (2008), a large pool of labeled turnover products was detected in mecillinam-treated samples (Figures 3B and 4). This observation confirms the conclusion that, rather than blocking PG synthesis by the Rod system like A22, mecillinam induced the degradation of nascent PG material produced by

In untreated cells, robust incorporation of label into the PG matrix was observed (Figure 4). Very little of the newly synthesized material was degraded as evidenced by the small amount of labeled anh MurNAc-peptides detected relative to the precursor pools (Figures 3B, 4, and S3). As reported by Uehara and Park (2008), we found the effects of A22 treatment differed dramatically from those induced by mecillinam. In cells treated with A22, incorporation of radiolabel into mature PG was significantly reduced as expected. This defect was accompanied by increased pools of the PG precursors mDAP and labeled UDP-MurNAc-pep₅ (Figures 3B, 4, and S3), suggestive of a late stage

the machine. Given that A22 suppressed the toxicity of mecillinam, we wondered if it would also block the apparent futile cycle of nascent PG synthesis and turnover induced by this beta-lactam. This indeed proved to be the case (Figures 3B and 4), thus linking the induction of nascent PG turnover with the lethal mode of action of mecillinam.

PBP inhibition by beta-lactams was previously found to decrease cellular levels of the final PG precursor, lipid II (Kohlrausch and Höltje, 1991). Based on in vitro PBP assays, it was recently suggested that the decline in lipid II concentration may result from the stimulation of glycan strand synthesis

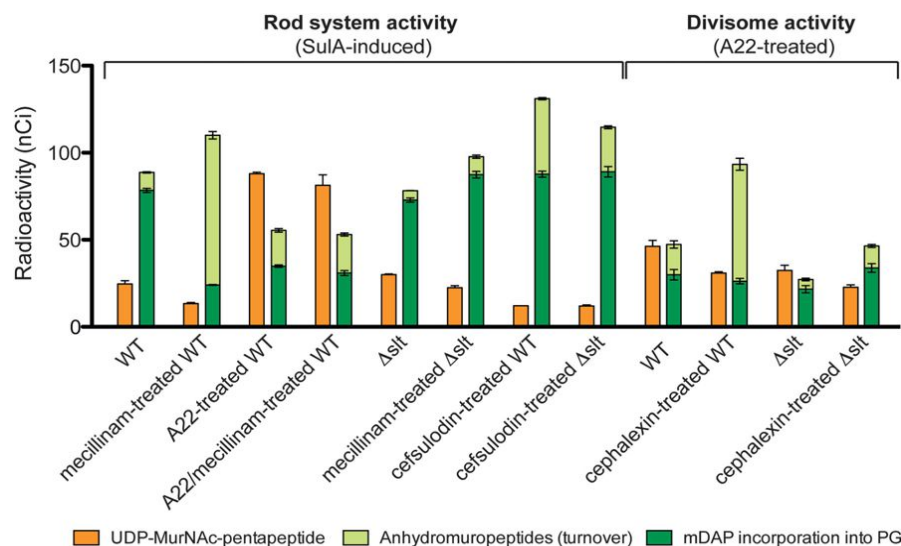


Figure 4. Quantification of PG Synthesis and Turnover following Beta-Lactam Treatment

Cells of TU278 [Δ lysA Δ ampD] or its Δ lts derivative (HC419) were grown, radiolabeled, and analyzed as in Figure 3. The amount of UDP-MurNAc-pentapeptide and total anhydromuropeptides produced were quantified from the area under the peaks in HPLC chromatograms. Radiolabel incorporation into PG was determined by quantifying the amount of label released from cells by lysozyme. Antibiotic concentrations used were: mecillinam (10 μ g/ml), A22 (10 μ g/ml), cephalixin (10 μ g/ml), and cefsulodin (100 μ g/ml). Results are the average of three independent experiments with the error bars representing the standard deviation. The drop in precursor levels between untreated and mecillinam-treated WT cells is significant ($p < 0.005$). See text for details.

when TP activity is blocked (Banzhaf et al., 2012). A dual mode of action for beta-lactams was thus proposed where TP inhibition and substrate depletion act synergistically to promote efficient bacterial killing (Banzhaf et al., 2012). Our results indicate that the futile-cycle of PG synthesis and turnover induced by beta-lactams is likely to be the main driver of the PG precursor depletion observed here and by others (Kohlrausch and Hölting, 1991; Uehara and Park, 2008) (Figure 3B and 4). To investigate the contribution of precursor depletion to the lethal activity of mecillinam, we generated strains that overexpress *uppS* or *murA*, which encode enzymes required for the synthesis of the lipid-carrier Und-P or UDP-MurNAc, respectively (Figure 3A). Importantly, both strains gained partial resistance to mecillinam (Figure S2D). Additionally, studies by D'Ari and colleagues (Vinella et al., 1993) showed that high-level overproduction of FtsZ (a 7 \times increase from pZAQ [Begg et al., 1998], as opposed to a 2 \times increase from pTB63 [Bernhardt and de Boer, 2005]), a condition where the divisome may more effectively compete with the Rod system for PG precursors, can partially suppress the lethal action of mecillinam. We conclude that mecillinam not only blocks the TP active site of PBP2, but by doing so, induces a futile cycle of cell wall synthesis and turnover by the Rod system that contributes to its bactericidal activity, at least in part, by depleting PG precursor pools.

Futile Cycle Induction Is a Common Property of Beta-Lactams

We investigated whether inhibiting PBPs other than PBP2 also stimulates the degradation of nascent PG. Cephalixin is a beta-lactam that preferentially targets PBP3, the cell division-specific bPBP (Spratt, 1975). To monitor the effect of cephalixin, we used a modified version of the labeling method described above. Instead of blocking division with SulA, cell cultures were treated with A22 to block Rod system activity prior to [3 H]-mDAP addition. This procedure allowed us to focus the analysis on divisome-mediated PG synthesis. In the absence of cephalixin, robust incorporation of label into PG by the divisome was detected, but with a higher basal level of PG turnover

than when the activity of the Rod system was monitored (Figures 4 and S4A). This observation is consistent with previous results (Uehara and Park, 2008), and most likely reflects the activity of the cell separation PG hydrolases. Importantly, just as with mecillinam, cephalixin treatment led to a dramatic increase in the degradation of nascent PG (Figures 4 and S4A).

The effect of blocking the TP activity of aPBPs was analyzed using cefsulodin, a beta-lactam that is specific for PBP1a and PBP1b (Curtis et al., 1979). Using labeling conditions identical to those for the mecillinam studies, we also observed a significant increase in nascent PG turnover upon cefsulodin treatment (Figures 4 and S4B). However, this increase was not accompanied by a corresponding drop in label incorporation into the PG matrix, presumably because PBP2 and other components of the Rod system remain functional when the TP activity of aPBPs is blocked. We conclude that the induction of nascent PG degradation is a common property of beta-lactams regardless of the growth process or class of PBP they target.

Slt Is Responsible for Nascent PG Turnover following Beta-Lactam Treatment

We reasoned that a genetic selection for mutants resistant to mecillinam might reveal the cell-wall-degrading enzyme responsible for the nascent PG turnover induced by beta-lactams. Notably, selections for mecillinam-resistance were performed many years ago by several laboratories (Spratt, 1975; Vinella et al., 1992; Wachi et al., 1987). This genetic analysis was instrumental in the initial identification of the genes encoding the Rod system (*mreBCD*, *pbpA*, and *rodA*). In retrospect, however, these selections were more complicated than originally appreciated because they were carried out in genetic backgrounds with wild-type FtsZ levels. Thus, based on our current understanding, these early mecillinam-resistance selections were actually demanding mutants with two changes: (1) those that increase FtsZ levels to render the Rod system nonessential, and (2) those that generate defects in the Rod system or other factors in order to alleviate the toxicity of the futile cycle of PG synthesis and degradation. Consistent with this view, the classic mecillinam-resistant mutants *mre-129*

and *mre-678* with lesions in the *mreBCD* operon (Wachi et al., 1987) were shown to produce elevated levels of FtsZ (Bendezú and de Boer, 2008). Therefore, to simplify the genetic analysis, we revisited the selections for mecillinam resistance in an FtsZ^{up} genetic background.

MG1655/pTB63 cells were mutagenized with a transposon and plated on LB agar supplemented with mecillinam. Similar to prior studies, two classes of resistant mutants were isolated: those that remained spherical in the absence of mecillinam and those that regained rod shape when the drug was removed. As expected, all of the permanently spherical mutants possessed transposon insertions that mapped to genes encoding components of the Rod system (data not shown). Mutants in the second class, however, retained rod-shaped without mecillinam and therefore must have a functional Rod system. We suspected that a subset of these mutants were likely to harbor insertions in genes important for the degradation of nascent PG following mecillinam treatment. One such mutant had an insertion in the *slt* gene encoding a periplasmic cell-wall-degrading enzyme called Soluble Lytic Transglycosylase (Slt).

Because LTs cleave glycan strands to generate the ^{anh}MurNAc-peptide products that accumulate following beta-lactam treatment, we focused on the *slt* mutant. To test whether Slt was responsible for beta-lactam-induced turnover of nascent PG, we performed the radiolabeling assay in a Δ *slt* strain. Without treatment, the loss of Slt function did not affect net PG synthesis by either the Rod system or the divisome, but the basal rate of turnover in each case was reduced (Figure 3B and 4). Significantly, following mecillinam or cephalexin treatment, the Δ *slt* strain displayed a dramatically reduced level of nascent PG turnover relative to the WT strain. Moreover, inactivation of Slt restored wild-type levels of PG incorporation into the matrix of mecillinam-treated cells (Figure 4). Slt also appeared to be important for PG turnover in cefsulodin-treated cells, but unlike the situation with drugs that inactivate bPBPs, a significant proportion of the turnover following aPBP inhibition was found to be Slt independent (Figure 4). Notably, the turnover products produced upon cefsulodin treatment had a much greater ^{anh}MurNAc-pentapeptide content (25%) than those observed following mecillinam treatment (4%) (Figures 3B and S4B). This observation suggests that uncrosslinked, nascent PG with a relatively high proportion of pentapeptides is formed by the cefsulodin-treated cells and that such material is subject to destruction by LTs other than Slt. Nevertheless, we conclude that Slt is likely to be the major enzyme responsible for nascent PG turnover following beta-lactam treatment.

Differential Effect of Slt Inactivation on Mecillinam Sensitivity Depending on FtsZ Levels

Curiously, in contrast to our results, Slt inactivation was previously found to result in mecillinam hypersensitivity rather than resistance (Templin et al., 1992). However, in a separate study, loss of Slt function was shown to increase mecillinam resistance in a *Salmonella typhimurium* mutant already partially resistant to the drug (Costa and Antón, 2006). Because the original observation of mecillinam hypersensitivity for a Δ *slt* mutant was made with otherwise wild-type cells, and our selection for resistance was performed with FtsZ^{up} cells, we suspected that the apparent

contradiction was related to cellular FtsZ concentration. Indeed, we confirmed that for cells lacking Slt and producing normal levels of FtsZ, the minimal inhibitory concentration (MIC) of mecillinam is almost ten times lower (0.04 μ g/ml) than that of wild-type cells (MIC = 0.3 μ g/ml). Conversely, as expected from our genetic analysis, FtsZ^{up} cells defective for Slt showed much higher resistance to mecillinam (MIC > 10 μ g/ml) than those with functional Slt (MIC = 0.3 μ g/ml). The change in mecillinam sensitivity of Δ *slt* cells relative to wild-type as a function of FtsZ levels was readily visualized on solid medium (Figure 5A). Importantly, Δ *slt* cells are not generally more sensitive to agents that target the Rod system; wild-type cells and a strain inactivated for Slt displayed the same A22 sensitivity (Figure S5A).

How can Slt inactivation result in both hypersensitivity and resistance to mecillinam? To investigate this conundrum, we compared the effect of mecillinam on the growth and morphology of wild-type and Δ *slt* cells with and without increased FtsZ production. Microscopic analysis following treatment with mecillinam for 3 hr revealed that cells lacking Slt undergo a morphological change at a much lower mecillinam concentration than wild-type cells. Moreover, the concentration of drug that leads to this shape defect was strongly correlated with the concentration that induces a growth defect in Δ *slt* cells (Figures 5 and S6). As with the growth phenotype, the morphology of Δ *slt* cells was only hypersensitive to mecillinam; wild-type and Δ *slt* cells were found to undergo shape changes at equivalent A22 concentrations (Figure S5). It has been well documented that the loss of rod shape results in cell division defects and poor growth unless FtsZ is overproduced (Bendezú and de Boer, 2008). We therefore conclude that the mecillinam hypersensitivity of Δ *slt* cells with normal FtsZ levels results from the morphological changes and consequent division inhibition induced by low doses of drug. On the contrary, when Δ *slt* cells have elevated FtsZ levels, they overcome the adverse effects of the shape alterations just as mutants with defects in the Rod system are suppressed by FtsZ overproduction.

Aberrant PG Crosslinking in Δ *slt* Cells Treated with Mecillinam

The shape change observed for Δ *slt* cells at low doses of mecillinam suggests that when uncrosslinked PG produced by beta-lactam-targeted synthetic complexes is not destroyed by Slt, it can be misincorporated into the PG matrix to alter morphology. To investigate this possibility, we analyzed the composition of purified cell walls (sacculi) isolated from wild-type or Δ *slt* cells following mecillinam treatment. In wild-type cells, the total amount of PG isolated following mecillinam treatment was reduced as expected, but the composition of the isolated material, including the relative degree of crosslinking, did not change dramatically (Figure 6A and Table S1). The mucopeptide profile of sacculi from untreated Δ *slt* cells was similar to that observed for wild-type PG. However, following mecillinam treatment, major alterations were observed (Figure 6B and Table S1). Strikingly, mucopeptides with PBP-generated crosslinks, those linked via the fourth residue of one peptide and the third residue of the other (a 4-3 crosslink), were almost completely depleted from mecillinam-treated Δ *slt* sacculi. This loss was accompanied by a corresponding increase in

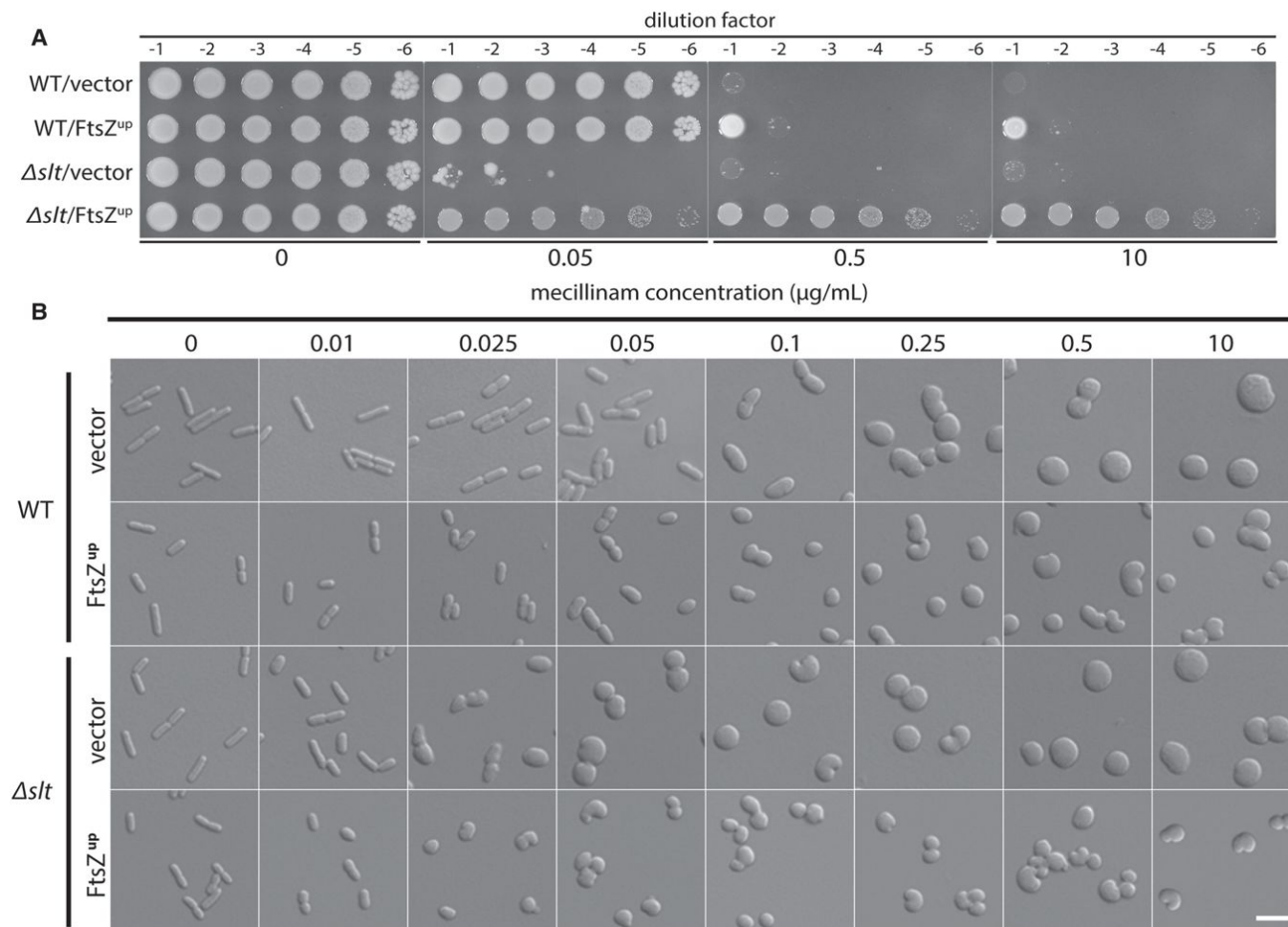


Figure 5. Effect of Mecillinam on the Growth and Morphology of Cells Defective for Slt

(A) Cells of MG1655 [WT] or HC408 [Δslt] harboring either pSC101 (vector) or pTB63 (FtsZ^{up}) were grown overnight, diluted, and spotted on agar containing the indicated concentration of mecillinam as described for Figure 2B.

(B) The same cells were grown to exponential phase and diluted to an OD₆₀₀ of 0.025 into LB medium containing the indicated concentration of mecillinam. Growth was continued for an additional 3 hr and the cells were fixed. The fixed cells were then imaged on agarose pads using DIC optics. Scale bar, 4 microns.

muropeptides with an alternative 3-3 (mDAP-mDAP) crosslink formed by L,D-transpeptidases. Thus, the failure to degrade uncrosslinked PG produced by beta-lactam-targeted synthetic machines indeed results in the aberrant incorporation of PG into the matrix. We conclude that the observed misincorporation is most likely responsible for the shape defects induced in Δslt cells treated with low levels of mecillinam. Furthermore, the accompanying drug hypersensitivity displayed by these mutants suggests that, in cells with wild-type levels of FtsZ, the shape changes induced by PG misincorporation are worse for viability than the consequences of the futile cycle, at least at low drug concentrations. Accordingly, the growth of Δslt cells with the vector control is more adversely affected by low doses of drug than the corresponding WT cells (Figure S6). At higher drug concentrations, however, functional Slt becomes detrimental, presumably due to the increased burden of the futile cycle as a higher proportion of Rod complexes become targeted (Figure S6).

DISCUSSION

A Common Theme for the Mode of Action of Bactericidal Drugs

Bactericidal drugs promote cell death while bacteriostatic agents merely stop bacterial growth. Antibiotics belonging to both general categories interfere with essential cellular pathways or enzymes. However, studies of the mode of action of clinically important drugs in the aminoglycoside and fluoroquinolone classes provide examples of how bactericidal agents typically do more than just inhibit their target enzymes. Aminoglycosides disable the proofreading capacity of the ribosome, causing the production of mistranslated proteins that are ultimately thought to cause irreversible membrane damage and death (Davis, 1987). Similarly, fluoroquinolones bind to DNA gyrase and topoisomerase IV, leading to the formation of stable drug-enzyme-DNA complexes that block DNA replication and promote the formation of double-strand breaks (Hooper, 2001). Thus, rather

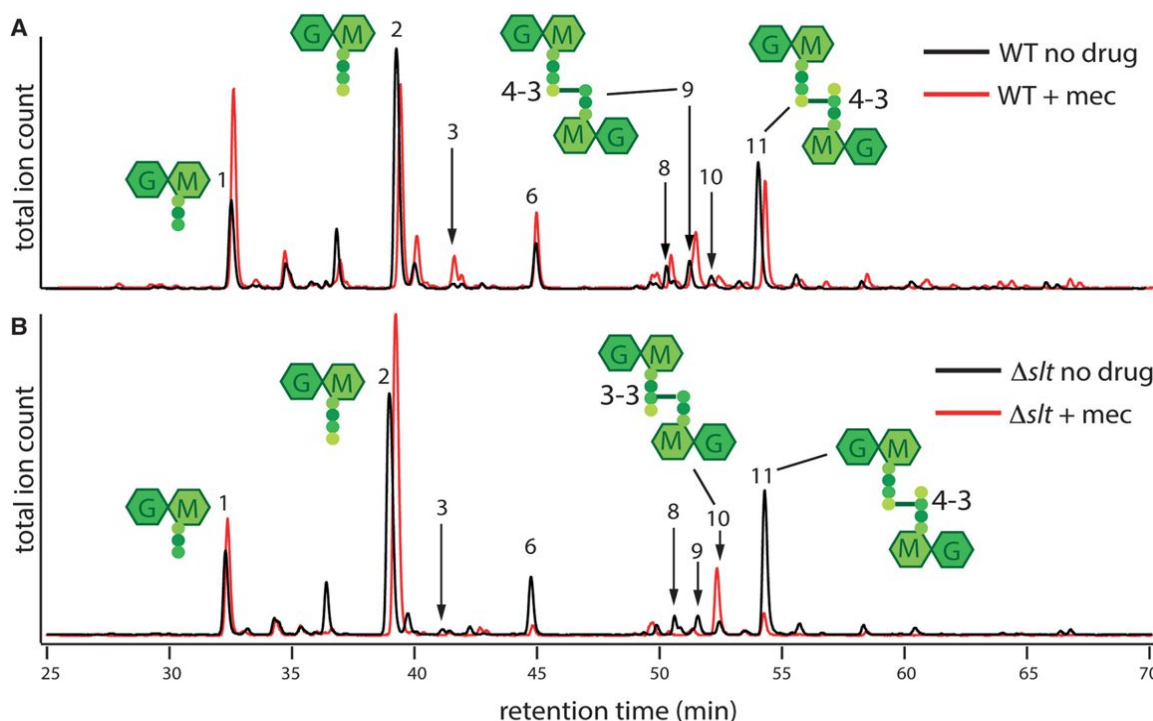


Figure 6. Muropeptide Composition of PG following Mecillinam Treatment

Cells of MG1655 [WT] (A) or HC408 [$\Delta s/t$] (B) were grown to an OD_{600} of approximately 0.5 and diluted 1:20 into fresh LB medium with or without mecillinam (10 $\mu\text{g/ml}$) as indicated. Growth was continued for an additional 3 hr and PG sacculi were prepared from cells of each culture. The purified PG was then digested with the muramidase mutanolysin and the resulting muropeptides were reduced and analyzed by LC/MS. Total ion count chromatograms are shown with the chromatograms of the mecillinam-treated samples off-set for clarity. Note that the scales of each chromatogram are not identical. They were scaled to show the relative muropeptide composition rather than the total amount of material. For example, the total peak area in the chromatogram from mecillinam-treated WT cells is one third that of the corresponding untreated sample even though an equivalent of twice the volume was injected. Schematics of several major muropeptide products are shown near their corresponding peaks with the numbers referring to the type of crosslink (4-3 v.s. 3-3). The identities for all labeled peaks and the quantification of their relative amounts are listed in Table S1. The results were reproducible over two biological replicates of each sample and three technical replicates for each biological replicate.

than simply inhibiting an essential activity, these drugs disrupt key functions of their target such that the activity of the crippled enzyme or multicomponent machine becomes toxic and reduces viability. Whether beta-lactams elicit a similar toxic malfunctioning of their target has remained unclear.

The widely accepted view of the mode of action of beta-lactams is that by inhibiting the TP activity of the PBPs, these drugs simply disrupt the balance between PG synthases and the action of PG hydrolases working to expand the cell wall matrix (Tomasz, 1979). Here, we show that beta-lactams derange the process in a more direct and insidious manner than this general framework suggests. Our results indicate that TP inhibition induces the turnover of nascent PG material produced by the targeted complex. In the case of PBP2 targeting by mecillinam, we clearly show that this turnover contributes significantly to the lethal action of the drug by inducing the depletion of cellular PG precursor pools such that even nontargeted PG synthetic complexes are likely to be adversely affected by beta-lactam treatment. Our findings thus connect the killing mechanisms of three important and widely used antibiotics (aminoglycosides, fluoroquinolones, and beta-lactams); they all stimulate a dominant-negative activity in their target pathway to induce systemic toxicity. Further-

more, all three drug classes typically function against multiple targets: aminoglycosides target ribosomes made from rRNA encoded by several genetic loci per genome (Davis, 1987), fluoroquinolones target gyrase and topoisomerase IV (Hooper, 2001), and beta-lactams typically target several different PBPs in a given organism (Curtis et al., 1979). It has been discussed previously that interfering with the activity of multiple cellular targets is an important property of monotherapies needed to reduce the incidence of mutational resistance (Silver, 2011). However, the potential benefit of such therapies inducing a dominant-negative activity in the affected pathway in addition to hitting multiple targets has been under appreciated. The combination of these properties likely enhances antibiotic effectiveness because, even if one of the targets is mutationally altered to block drug binding, the sensitivity and lethal malfunctioning of the other targets will be dominant over the lone resistant allele. Thus, the principles learned from this and other studies of the mode of action of our classic antibacterial therapies indicate that, rather than searching for simple inhibitors of essential enzymes as has been common practice, antibiotic discovery efforts should ideally be seeking new molecules that induce a lethal malfunctioning of multiple cellular targets.

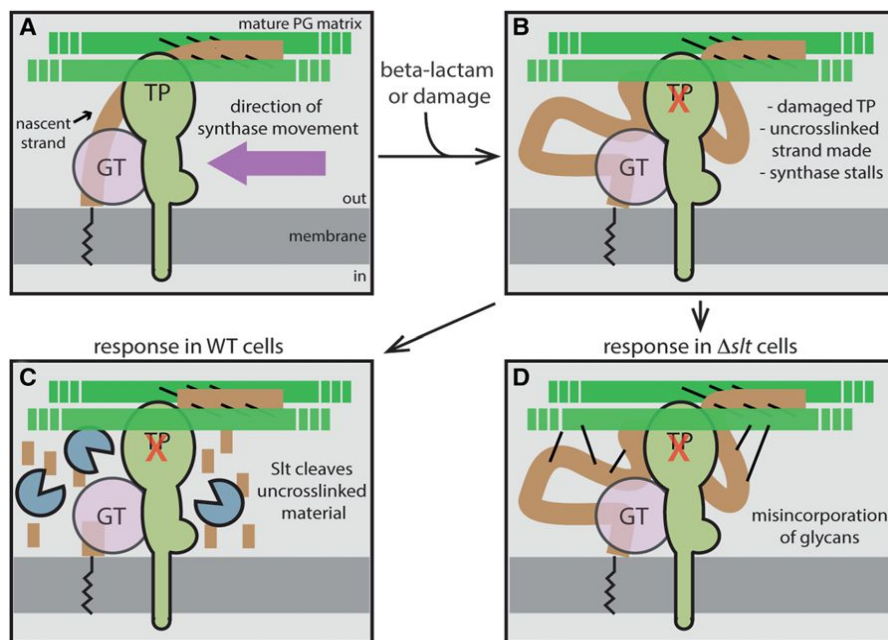


Figure 7. Role of Slt in Maintaining Coordination between GT and TP Activity during PG Biogenesis

Shown is a schematic highlighting the role of Slt in PG biogenesis. (A) The normal synthetic process with properly coupled GT and TP activities. The new glycan strands (brown) are polymerized from lipid-linked precursors (black zig-zag line) by an enzyme with GT activity (pink) that may either be the canonical domain of an aPBP or as an yet unidentified GT enzyme. The newly polymerized glycans are rapidly crosslinked into the mature matrix (green) by an associated TP activity shown here as a bPBP, but it could also be the TP domain of an aPBP. Other components of the putative synthetic complex including cytoskeletal elements were omitted for simplicity. When the TP active site is damaged or targeted by a beta-lactam (B), GT activity continues to produce glycan chains that are no longer cross-linked into the matrix. In WT cells (C), such strands are targeted for destruction by Slt. Conversely, in cells inactivated for Slt (D), the uncrosslinked glycans produced by the damaged/targeted machinery are not degraded, but are instead aberrantly incorporated into the matrix by an alternative crosslinking enzyme. In cells with normal FtsZ levels, the resulting morphological changes are lethal, leading to beta-lactam hypersensitivity of Slt-defective cells.

A Potential Role for Slt as a Quality Control Enzyme in PG Biogenesis

In addition to revealing novel features of the beta-lactam killing mechanism, our findings also shed significant new light on the process of PG biogenesis. Pulse-labeling studies of PG assembly indicate that nascent PG material is rapidly crosslinked into the mature matrix (Burman and Park, 1983; Glauner and Hölting, 1990), indicating a tight coupling between GT and TP activities. Such an orderly insertion mechanism is likely critical for the maintenance of proper cell shape and integrity. Here, we show that the inhibition of TP activity by beta-lactams functions to uncouple glycan polymerization from crosslinking, leading to the generation of uncrosslinked glycan strands (Figure 7). Our results indicate that this nascent material is rapidly degraded in an Slt-dependent manner, setting up a futile cycle of synthesis and degradation (Figure 7). Mutants defective for other PG cleaving enzymes were not isolated in selections for mecillinam resistance. Additionally, nascent PG turnover was largely eliminated in the absence of Slt. We therefore conclude that Slt is likely to be directly responsible for destroying the uncrosslinked material produced by the beta-lactam-targeted synthetic machinery.

How Slt is activated to degrade uncrosslinked PG is not known. Importantly, affinity chromatography studies identified a potential interaction of Slt with PBPs 1b, 2, and 3 in *E. coli* (von Rechenberg et al., 1996; Romeis and Hölting, 1994), suggesting that the turnover we observe may be stimulated by association with the beta-lactam-targeted synthases. Additionally, Slt is known to have a doughnut-like structure with the active site oriented toward a central hole wide enough to accommodate only a single uncrosslinked glycan chain (Thunnissen et al., 1994). Thus, the enzyme may have a strong substrate preference for uncrosslinked glycan chains, a possibility that is sup-

ported by the observation that pretreatment of purified cell wall sacculi with an endopeptidase stimulates the release of soluble material by Slt (Romeis and Hölting, 1994). This apparent substrate preference may contribute to its role in the futile cycle in addition to, or independent of, an interaction with the PBPs. Notably, Slt functions primarily as an exoglycosidase (Beachey et al., 1981) with some detectable endoglycosidase activity (Lee et al., 2013). Although it has been argued that Slt degrades glycan chains from an anhydro-MurNAc end (van Asselt et al., 1999), the structural data are also consistent with Slt processing glycan strands from the nonreducing end as has been demonstrated biochemically. We therefore envision that Slt may either recognize an existing, free nonreducing end of an uncrosslinked glycan, or, in the case of partially crosslinked material, it could create such an end with its weak endoglycosidase activity, followed by the degradation of the remaining uncrosslinked glycan chain.

The physiological role of Slt has remained mysterious for some time. It has been implicated in the turnover of mature cell wall material and cell separation (Heidrich et al., 2002; Kraft et al., 1999). However, these functions have been ascribed to Slt based on the phenotypes of mutants lacking multiple PG processing enzymes, suggesting that they represent secondary activities. Clues to the likely primary physiological function of Slt were provided by our analysis of the response of Δslt cells to mecillinam treatment. In the absence of Slt, PG produced by mecillinam-targeted synthetic complexes was found to be aberrantly incorporated into the PG matrix via the formation of alternative (3-3) crosslinks by non-PBP transpeptidases. Additionally, the morphology of Δslt cells was dramatically altered following exposure to low doses of mecillinam, while the same treatment had little effect on the shape of wild-type cells. Because these drug

concentrations are well below the IC_{50} for PBP2 (Curtis et al., 1979), only a fraction of the cellular PBP2 pool is predicted to be inhibited in cells treated with this level of mecillinam. Thus, the morphological changes displayed by $\Delta s/t$ cells exposed to low concentrations of mecillinam are unlikely to be due to general inhibition of the Rod system. Instead, we conclude that misincorporation of uncrosslinked PG produced by mecillinam-targeted machines is likely to be responsible for the shape change of the mutant, most likely because the aberrant material creates a local defect in the matrix that interferes with the proper assembly or remodeling of PG by other (untargeted) machines. By extension, we infer that a major function of Slt is to prevent the stable accumulation of uncrosslinked glycan strands so that they are not aberrantly crosslinked into the mature matrix. Such an activity is consistent with a novel “quality control” or “repair” function that helps maintain cell shape and integrity.

In our experimental system, low drug concentration most likely serves to mimic environmental damage to the synthetic machinery or other natural conditions resulting in the uncoupling of GT and TP activities. For example, uncrosslinked glycan strands might accumulate if synthetic complexes enter an area of the PG matrix with few peptides available to accept crosslinks. In these cases, we envision that glycan chain degradation by Slt might facilitate the exchange of a damaged TP for a functional one, and/or allow the release of stalled complexes tethered to a partially crosslinked glycan strand so that they can restart synthesis at other locations. In support of the latter possibility, recent studies indicate that bPBPs are likely to be dynamically associated with the synthetic machinery (Lee et al., 2014).

An additional phenotype consistent with a quality control function for Slt is the rapid lysis displayed by $\Delta s/t$ cells treated with cephalaxin (Templin et al., 1992). PBP3 inactivation normally results in a division block and the formation of filamentous cells that lyse many mass doublings after cephalaxin addition (Spratt, 1975). However, cells defective for Slt lyse rapidly following cephalaxin treatment (Templin et al., 1992). Rapid lysis under these conditions is known to require the activity of the cell separation amidases (Heidrich et al., 2001). By analogy with mecillinam-treated cells, addition of cephalaxin to $\Delta s/t$ cells likely results in the production of stable uncrosslinked glycan strands that can be misincorporated by alternative crosslinking enzymes. Lysis may result because this misincorporated material mimics productive PG synthesis by the divisome such that the cell separation enzymes are activated to process the malformed material resulting in a localized breach in the PG. Thus, in addition to shape maintenance, the proposed quality control function of Slt may also enforce the coupling of GT and TP activities to prevent the misactivation of PG hydrolases by damaged PG synthases. Taking this line of reasoning a step further, it is interesting to speculate that it may be the failure of Slt to process certain types of glycan strands (e.g., those with high pentapeptide content) that underlies the ability of some beta-lactam drugs to induce rapid cell lysis via the misactivation of autolytic PG hydrolases.

Slt, Beta-Lactamase Induction, and Potential Combination Therapies

Many gram-negative bacteria encode a beta-lactamase (AmpC) produced in response to beta-lactam treatment (Fisher and Mo-

bashery, 2014). In some organisms, expression of the *ampC* gene is controlled by the transcriptional regulator AmpR (Jacobs et al., 1997). Mutants defective for the AmpD amidase involved in cell wall recycling were found to overexpress *ampC* even in the absence of beta-lactams (Jacobs et al., 1994). This finding along with supporting in vitro transcription experiments has led to a model in which beta-lactam treatment leads to the accumulation of ^{anh}MurNAc-peptide turnover products in the cytoplasm that serve as a signal for *ampC* induction (Fisher and Mobashery, 2014; Jacobs et al., 1997). It has been widely assumed that these recycling intermediates increase in concentration because beta-lactams induce general cell wall damage. Here, we have shown that targeting PG synthetic machines with beta-lactams causes nascent PG to be degraded by Slt, generating high levels of ^{anh}MurNAc-peptides. Thus, rather than generic insults to the PG matrix, the *ampR-ampC* regulatory system is perfectly tuned to detect the futile cycle of PG synthesis and degradation induced upon TP inhibition. Importantly, based on the critical role observed for Slt in *ampC* induction (Kraft et al., 1999) and the hypersensitivity of Slt-defective mutants to beta-lactams (Templin et al., 1992), it has been proposed that Slt inhibitors may be useful in combination therapies with beta-lactams (Kraft et al., 1999). Our results provide a mechanistic picture of what such combinations would likely accomplish. The downside of blocking Slt activity simultaneously with beta-lactam treatment would be loss of the toxic effects of the futile cycle and subsequent precursor depletion. However, what would be gained is an alternative malfunction of PG biogenesis caused by the accumulation of uncrosslinked glycan strands that can cause lysis or lethal shape changes as described above. Thus, inhibiting Slt function should indeed be an effective way to stymie beta-lactamase induction systems in gram-negative pathogens while still allowing beta-lactams to derange the PG biogenesis machinery to promote cell death.

CONCLUSION

Beta-lactams are arguably one of the most important therapeutic classes in the history of medicine. In this report, we present data that changes the view of how these drugs work. They are not simple inhibitors of cell wall synthesis as is commonly believed. Instead, they derange the activity of the PG biogenesis machinery to induce a futile cycle of PG synthesis and degradation that depletes cellular resources. While investigating the nature of the futile cycle, we uncovered what appears to be a new role for PG-cleaving enzymes as quality control factors that intervene when problems arise during PG synthesis. Overall, these new insights provide a mechanistic foundation that will not only help us better understand the process of cell wall biogenesis, but also how best to interfere with it for the development of novel antibacterial treatments.

EXPERIMENTAL PROCEDURES

Media, Bacterial Strains, Plasmids, and Culture Conditions

Cells were grown in LB or minimal M9 medium supplemented with 0.2% caseamino acids and carbon source (0.4% glycerol or 0.2% maltose) as indicated. The bacterial strains and plasmids used in this study are listed in Tables S2 and S3, respectively.

Measurement of the Turnover of Newly Synthesized PG

The effect of beta-lactams and A22 on the turnover of nascent PG was monitored by using $\Delta lysA \Delta ampD$ strains essentially as described previously (Uehara and Park, 2008). Cells of the relevant strains were grown overnight in M9-glycerol medium with 0.2% casamino acids. The resulting cultures were then diluted to an OD₆₀₀ of 0.04 in the same medium and grown to an OD₆₀₀ between 0.3–0.4 at 30°C. For experiments measuring PG synthesis and turnover by the Rod complex, *sulA* expression was induced from a chromosomally integrated plasmid (pHC739) for 30 min before drug treatment. The cultures were then treated with the indicated beta-lactams and/or A22 for 5 min after adjusting the culture OD₆₀₀ to 0.3. Following drug treatment, [³H]-mDAP (1 μCi) was added to 1 ml of each culture for 10 min (1/10 of the doubling time) to label the newly synthesized PG and its turnover products. After the labeling, cells were pelleted, resuspended in 0.7 ml water, and heated at 90°C for 30 min to extract water-soluble compounds. After the hot water extraction, insoluble material was pelleted by ultracentrifugation. The resulting supernatant was then removed, lyophilized, and resuspended in 0.1% formic acid for HPLC analysis. To determine [³H]-mDAP incorporated into the PG matrix, the pellet fraction was washed and treated with lysozyme. The suspensions were incubated overnight at 37°C. Insoluble material was then pelleted by centrifugation, and the resulting supernatant was subjected to scintillation counting. Similar procedures were used for measuring PG synthesis and turnover by the divisome following cephalixin treatment except that Rod system activity was suppressed by treating the cultures with A22. See Supplemental Information for a detailed protocol.

Muropeptide Analysis of PG Sacculi Isolated from Drug-Treated Cells

Cultures of MG1655 and its Δslt derivative, HC408, were diluted in 1 l of LB or LB containing 10 μg/ml mecillinam to OD₆₀₀ of 0.025 and incubated at 30°C for 3 hr with shaking. The OD₆₀₀ of each culture reached approximately 0.6 at the end of the growth period. PG was then purified from each culture and analyzed by LC/MS as described in Supplemental Information.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, six figures, and three tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2014.11.017>.

AUTHOR CONTRIBUTIONS

H.C. designed, performed, and interpreted all experiments. T.U. designed, performed and interpreted critical foundational studies for this work. T.G.B. designed and interpreted the experiments. The paper was written and edited by T.G.B., H.C., and T.U.

ACKNOWLEDGMENTS

The authors would like to thank Christina Montero and Alex Godfrey for help with mutant characterization and members of the Bernhardt, Rudner, and Walker laboratories for helpful discussions. We also thank Tom Dougherty, Mary-Jane Tsang, Anastasiya Yakhnina, and David Rudner for comments on the manuscript. Special thanks to Suzanne Walker and her laboratory their help with LC/MS. This work was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (R01-AI099144 and U19-AI109764).

Received: June 1, 2014

Revised: August 25, 2014

Accepted: November 10, 2014

Published: December 4, 2014

REFERENCES

Banzhaf, M., van den Berg van Saparoea, B., Terrak, M., Fraipont, C., Egan, A., Philippe, J., Zapun, A., Breukink, E., Nguyen-Distèche, M., den Blaauwen, T.,

and Vollmer, W. (2012). Cooperativity of peptidoglycan synthases active in bacterial cell elongation. *Mol. Microbiol.* **85**, 179–194.

Beachey, E.H., Keck, W., de Pedro, M.A., and Schwarz, U. (1981). Exoenzymatic activity of transglycosylase isolated from *Escherichia coli*. *Eur. J. Biochem.* **116**, 355–358.

Begg, K., Nikolaichik, Y., Crossland, N., and Donachie, W.D. (1998). Roles of FtsA and FtsZ in activation of division sites. *J. Bacteriol.* **180**, 881–884.

Bendezú, F.O., and de Boer, P.A.J. (2008). Conditional lethality, division defects, membrane involution, and endocytosis in *mre* and *mrd* shape mutants of *Escherichia coli*. *J. Bacteriol.* **190**, 1792–1811.

Bernhardt, T.G., and de Boer, P.A.J. (2005). SlmA, a nucleoid-associated, FtsZ binding protein required for blocking septal ring assembly over Chromosomes in *E. coli*. *Mol. Cell* **18**, 555–564.

Burman, L.G., and Park, J.T. (1983). Changes in the composition of *Escherichia coli* murein as it ages during exponential growth. *J. Bacteriol.* **155**, 447–453.

Chung, H.S., Yao, Z., Goehring, N.W., Kishony, R., Beckwith, J., and Kahne, D. (2009). Rapid beta-lactam-induced lysis requires successful assembly of the cell division machinery. *Proc. Natl. Acad. Sci. USA* **106**, 21872–21877.

Costa, C.S., and Antón, D.N. (2006). High-level resistance to mecillinam produced by inactivation of soluble lytic transglycosylase in *Salmonella enterica* serovar Typhimurium. *FEMS Microbiol. Lett.* **256**, 311–317.

Curtis, N.A., Orr, D., Ross, G.W., and Boulton, M.G. (1979). Affinities of penicillins and cephalosporins for the penicillin-binding proteins of *Escherichia coli* K-12 and their antibacterial activity. *Antimicrob. Agents Chemother.* **16**, 533–539.

Davis, B.D. (1987). Mechanism of bactericidal action of aminoglycosides. *Microbiol. Rev.* **51**, 341–350.

de Boer, P.A.J. (2010). Advances in understanding *E. coli* cell fission. *Curr. Opin. Microbiol.* **13**, 730–737.

Fisher, J.F., and Mobashery, S. (2014). The sentinel role of peptidoglycan recycling in the β-lactam resistance of the Gram-negative Enterobacteriaceae and *Pseudomonas aeruginosa*. *Bioorg. Chem.* **56**, 41–48.

Gitai, Z., Dye, N.A., Reisenauer, A., Wachi, M., and Shapiro, L. (2005). MreB actin-mediated segregation of a specific region of a bacterial chromosome. *Cell* **120**, 329–341.

Glauner, B., and Höltje, J.V. (1990). Growth pattern of the murein sacculus of *Escherichia coli*. *J. Biol. Chem.* **265**, 18988–18996.

Heidrich, C., Templin, M.F., Ursinus, A., Merdanovic, M., Berger, J., Schwarz, H., de Pedro, M.A., and Höltje, J.V. (2001). Involvement of N-acetylmuramyl-L-alanine amidases in cell separation and antibiotic-induced autolysis of *Escherichia coli*. *Mol. Microbiol.* **41**, 167–178.

Heidrich, C., Ursinus, A., Berger, J., Schwarz, H., and Höltje, J.-V. (2002). Effects of multiple deletions of murein hydrolases on viability, septum cleavage, and sensitivity to large toxic molecules in *Escherichia coli*. *J. Bacteriol.* **184**, 6093–6099.

Hooper, D.C. (2001). Mechanisms of action of antimicrobials: focus on fluoroquinolones. *Clin. Infect. Dis.* **32** (Suppl 1), S9–S15.

Jacobs, C., Huang, L.J., Bartowsky, E., Normark, S., and Park, J.T. (1994). Bacterial cell wall recycling provides cytosolic muropeptides as effectors for beta-lactamase induction. *EMBO J.* **13**, 4684–4694.

Jacobs, C., Frère, J.-M., and Normark, S. (1997). Cytosolic intermediates for cell wall biosynthesis and degradation control inducible β-lactam resistance in gram-negative bacteria. *Cell* **88**, 823–832.

Kohlrausch, U., and Höltje, J.V. (1991). Analysis of murein and murein precursors during antibiotic-induced lysis of *Escherichia coli*. *J. Bacteriol.* **173**, 3425–3431.

Kraft, A.R., Prabhu, J., Ursinus, A., and Höltje, J.V. (1999). Interference with murein turnover has no effect on growth but reduces beta-lactamase induction in *Escherichia coli*. *J. Bacteriol.* **181**, 7192–7198.

- Kruse, T., Bork-Jensen, J., and Gerdes, K. (2005). The morphogenetic MreBCD proteins of *Escherichia coli* form an essential membrane-bound complex. *Mol. Microbiol.* **55**, 78–89.
- Lee, M., Heseck, D., Llarrull, L.I., Lastochkin, E., Pi, H., Boggess, B., and Mobashery, S. (2013). Reactions of all *Escherichia coli* lytic transglycosylases with bacterial cell wall. *J. Am. Chem. Soc.* **135**, 3311–3314.
- Lee, T.K., Tropini, C., Hsin, J., Desmarais, S.M., Ursell, T.S., Gong, E., Gitai, Z., Monds, R.D., and Huang, K.C. (2014). A dynamically assembled cell wall synthesis machinery buffers cell growth. *Proc. Natl. Acad. Sci. USA* **111**, 4554–4559.
- Moreillon, P., Markiewicz, Z., Nachman, S., and Tomasz, A. (1990). Two bactericidal targets for penicillin in pneumococci: autolysis-dependent and autolysis-independent killing mechanisms. *Antimicrob. Agents Chemother.* **34**, 33–39.
- Park, J.T., and Strominger, J.L. (1957). Mode of action of penicillin. *Science* **125**, 99–101.
- Romeis, T., and Höltje, J.V. (1994). Specific interaction of penicillin-binding proteins 3 and 7/8 with soluble lytic transglycosylase in *Escherichia coli*. *J. Biol. Chem.* **269**, 21603–21607.
- Sauvage, E., Kerff, F., Terrak, M., Ayala, J.A., and Charlier, P. (2008). The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiol. Rev.* **32**, 234–258.
- Schwarz, U., Asmus, A., and Frank, H. (1969). Autolytic enzymes and cell division of *Escherichia coli*. *J. Mol. Biol.* **41**, 419–429.
- Silver, L.L. (2011). Challenges of antibacterial discovery. *Clin. Microbiol. Rev.* **24**, 71–109.
- Spratt, B.G. (1975). Distinct penicillin binding proteins involved in the division, elongation, and shape of *Escherichia coli* K12. *Proc. Natl. Acad. Sci. USA* **72**, 2999–3003.
- Templin, M.F., Edwards, D.H., and Höltje, J.V. (1992). A murein hydrolase is the specific target of bulgecin in *Escherichia coli*. *J. Biol. Chem.* **267**, 20039–20043.
- Thunnissen, A.M., Dijkstra, A.J., Kalk, K.H., Rozeboom, H.J., Engel, H., Keck, W., and Dijkstra, B.W. (1994). Doughnut-shaped structure of a bacterial muramidase revealed by X-ray crystallography. *Nature* **367**, 750–753.
- Tipper, D.J., and Strominger, J.L. (1965). Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. *Proc. Natl. Acad. Sci. USA* **54**, 1133–1141.
- Tomasz, A. (1979). The mechanism of the irreversible antimicrobial effects of penicillins: how the beta-lactam antibiotics kill and lyse bacteria. *Annu. Rev. Microbiol.* **33**, 113–137.
- Tomasz, A., and Waks, S. (1975). Mechanism of action of penicillin: triggering of the pneumococcal autolytic enzyme by inhibitors of cell wall synthesis. *Proc. Natl. Acad. Sci. USA* **72**, 4162–4166.
- Tomasz, A., Albino, A., and Zanati, E. (1970). Multiple antibiotic resistance in a bacterium with suppressed autolytic system. *Nature* **227**, 138–140.
- Typas, A., Banzhaf, M., Gross, C.A., and Vollmer, W. (2012). From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat. Rev. Microbiol.* **10**, 123–136.
- Uehara, T., and Bernhardt, T.G. (2011). More than just lysins: peptidoglycan hydrolases tailor the cell wall. *Curr. Opin. Microbiol.* **14**, 698–703.
- Uehara, T., and Park, J.T. (2008). Growth of *Escherichia coli*: significance of peptidoglycan degradation during elongation and septation. *J. Bacteriol.* **190**, 3914–3922.
- Uehara, T., Dinh, T., and Bernhardt, T.G. (2009). LytM-domain factors are required for daughter cell separation and rapid ampicillin-induced lysis in *Escherichia coli*. *J. Bacteriol.* **191**, 5094–5107.
- van Asselt, E.J., Thunnissen, A.M., and Dijkstra, B.W. (1999). High resolution crystal structures of the *Escherichia coli* lytic transglycosylase Slt70 and its complex with a peptidoglycan fragment. *J. Mol. Biol.* **297**, 877–898.
- Vinella, D., D'Ari, R., Jaffé, A., and Boulloc, P. (1992). Penicillin binding protein 2 is dispensable in *Escherichia coli* when ppGpp synthesis is induced. *EMBO J.* **11**, 1493–1501.
- Vinella, D., Joseleau-Petit, D., Thévenet, D., Boulloc, P., and D'Ari, R. (1993). Penicillin-binding protein 2 inactivation in *Escherichia coli* results in cell division inhibition, which is relieved by FtsZ overexpression. *J. Bacteriol.* **175**, 6704–6710.
- von Rechenberg, M., Ursinus, A., and Höltje, J.V. (1996). Affinity chromatography as a means to study multienzyme complexes involved in murein synthesis. *Microb. Drug Resist.* **2**, 155–157.
- Wachi, M., Doi, M., Tamaki, S., Park, W., Nakajima-Iijima, S., and Matsushashi, M. (1987). Mutant isolation and molecular cloning of mre genes, which determine cell shape, sensitivity to mecillinam, and amount of penicillin-binding proteins in *Escherichia coli*. *J. Bacteriol.* **169**, 4935–4940.

Tissue-Resident Macrophage Enhancer Landscapes Are Shaped by the Local Microenvironment

Yonit Lavin,^{1,3} Deborah Winter,^{2,3} Ronnie Blecher-Gonen,^{2,3} Eyal David,² Hadas Keren-Shaul,² Miriam Merad,^{1,4,*} Steffen Jung,^{2,4,*} and Ido Amit^{2,4,*}

¹Department of Oncological Sciences, Immunology Institute and the Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

²Department of Immunology, Weizmann Institute of Science, Rehovot 76100, Israel

³Co-first author

⁴Co-senior author

*Correspondence: miriam.merad@mssm.edu (M.M.), s.jung@weizmann.ac.il (S.J.), ido.amit@weizmann.ac.il (I.A.)

<http://dx.doi.org/10.1016/j.cell.2014.11.018>

SUMMARY

Macrophages are critical for innate immune defense and also control organ homeostasis in a tissue-specific manner. They provide a fitting model to study the impact of ontogeny and microenvironment on chromatin state and whether chromatin modifications contribute to macrophage identity. Here, we profile the dynamics of four histone modifications across seven tissue-resident macrophage populations. We identify 12,743 macrophage-specific enhancers and establish that tissue-resident macrophages have distinct enhancer landscapes beyond what can be explained by developmental origin. Combining our enhancer catalog with gene expression profiles and open chromatin regions, we show that a combination of tissue- and lineage-specific transcription factors form the regulatory networks controlling chromatin specification in tissue-resident macrophages. The environment is capable of shaping the chromatin landscape of transplanted bone marrow precursors, and even differentiated macrophages can be reprogramed when transferred into a new microenvironment. These results provide a comprehensive view of macrophage regulatory landscape and highlight the importance of the microenvironment, along with pioneer factors in orchestrating identity and plasticity.

INTRODUCTION

Macrophages are hematopoietic cells of the myeloid lineage that are specialized in phagocytosis and respond to diverse environmental signals (Epelman et al., 2014; Ginhoux and Jung, 2014; Lavin and Merad, 2013; van Furth et al., 1972). They actively maintain steady state by secreting and responding to cytokines and chemokines (Mortha et al., 2014; Zigmond et al., 2014). In addition, tissue-resident macrophages play important homeo-

static roles, depending on the tissue in which they reside. Microglia, the brain-resident macrophages, prune synapses during development (Paolicelli et al., 2011; Schafer et al., 2012). Spleen red pulp macrophages phagocytose erythrocytes and recycle heme to maintain iron homeostasis (Chow et al., 2013; Kohyama et al., 2009). Peritoneal cavity macrophages regulate the production of gut immunoglobulin (Ig) A by interacting with peritoneal B1 cells (Okabe and Medzhitov, 2014). These studies, among others, highlight the plasticity of macrophages and their specialization to fulfill tissue-specific functions.

Recent studies have demonstrated that most tissues are populated early during fetal development by macrophages that subsequently maintain themselves, independently of adult hematopoiesis, through longevity and limited self-renewal (Ginhoux et al., 2010; Hashimoto et al., 2013; Schulz et al., 2012; Yona et al., 2013). Thus, most macrophages, although sharing a common lineage, take residence in tissues early during embryogenesis, and the respective macrophage compartments develop locally and independently from each other. A notable exception from this scheme is macrophages residing in the intestine, as these cells are constantly replenished from monocytes even in steady state (Bain et al., 2014; Bogunovic et al., 2009; Varol et al., 2009). Thus, distinct ontogeny is one defining feature of macrophages, but it is unclear to what extent it shapes macrophage identity.

Emerging evidence indicates that environmental factors influence the specialization of tissue-resident macrophages. Heme has been shown to induce Spi-c, a transcription factor (TF) important for red pulp macrophage development (Haldar et al., 2014; Kohyama et al., 2009). Retinoic acid (RA) stimulates Gata6 expression and thereby contributes to the regulatory program of peritoneal macrophages (Okabe and Medzhitov, 2014). Finally, TGF- β was shown to regulate a microglia expression program through Smad TFs (Abutbul et al., 2012; Butovsky et al., 2014). These limited reports provide evidence that environment can govern the expression of tissue-specific macrophage signatures.

Epigenetic modification is one conduit through which ontogeny and environment can influence the development of macrophage identities. The chromatin landscape, among other epigenomic features of a differentiated cell type, reflects both

its developmental origin, as well as its future potential (Lara-Astiaso et al., 2014; Stergachis et al., 2013; Winter and Amit, 2014). Nucleosomes are the fundamental unit of chromatin consisting of ~147 bases of DNA wrapped around a histone core. Nucleosome-depleted regions, known as “open chromatin,” contain regulatory elements—such as promoters and enhancers—that play a critical role in gene regulation (Gross and Garrard, 1988). Changes enacted by chromatin remodelers, such as nucleosome eviction or insertion, as well as the addition or deletion of histone modifications, have been linked to changes in the expression of nearby genes (Cirillo et al., 2002; Felsenfeld and Groudine, 2003). Many regulatory modifications are ubiquitous, but variations on a global scale generate the distinct chromatin landscape observed between cell types (Ernst et al., 2011; Heintzman et al., 2009).

During development, pioneer TFs initiate chromatin accessibility to allow the binding of additional TFs (Cirillo et al., 2002; Garber et al., 2012). PU.1 has been implicated as a pioneering factor throughout hematopoietic development, especially in the myeloid lineage. In macrophages, PU.1 occupies most enhancers, where it is necessary for the maintenance of methylation on the fourth lysine of the H3 subunit (H3K4me1) (Ghisletti et al., 2010; Heinz et al., 2010). The cobinding of PU.1 with lineage-specific TFs orchestrates cell-type specificity by regulating expression and establishing the chromatin landscape (Heinz et al., 2010; Laslo et al., 2006). Cell-type-specific responses to stimuli are largely coordinated through activation by stimulus-triggered TFs that frequently bind to previously occupied “poised” enhancers (Garber et al., 2012; Ostuni et al., 2013). Poised enhancers may reflect past activity and persist throughout development or arise during lineage specification (Lara-Astiaso et al., 2014). Active enhancers mark the current state of a cell and can be distinguished by the presence of acetyl groups on the histone tails, particularly H3K27ac (Creyghton et al., 2010; Heintzman et al., 2007). Tissue-resident macrophages provide a fitting model for examining how chromatin is programmed through development to allow for plasticity within a cell type to specify tissue-specific functions.

Here, we profile the expression and chromatin landscape of seven populations of mouse macrophages isolated from distinct tissues to determine the contributions of ontogeny and environment to tissue-resident macrophage identity. Based on the distribution of histone modifications, we map the regulatory elements of tissue-resident macrophages, including promoters, active enhancers, and poised enhancers. We compare candidate enhancers marked with H3K4me1 in different macrophage populations with those found in monocytes or neutrophils. Through analyzing the distinct enhancer landscape of tissue-resident macrophages, we assess the impact of developmental origin and local microenvironment and identify several potential regulators from TF-binding motifs within these regions. We further assess the essential contribution of tissue environment by replacing endogenous macrophages with adult bone-marrow-derived cells and transferring differentiated macrophages into a new tissue microenvironment. Collectively, our results indicate that, aside from ontogeny, the environment plays a critical role in shaping the unique identity and function of tissue-resident macrophages through the regulation of TFs.

RESULTS

Genome-wide Assays to Identify Regulatory Elements in Macrophages

To elucidate the transcriptional and epigenomic networks in tissue-resident macrophages, we performed an array of complementary genome-wide assays on at least two biological replicates followed by high-throughput sequencing: RNA-seq, chromatin immunoprecipitation (ChIP-seq), and an assay for transposase-accessible chromatin (ATAC-seq) (Figure 1A). We purified macrophages, monocytes, and neutrophils from fresh mouse tissues using fluorescence-activated cell sorting (FACS) (Experimental Procedures). Global gene expression profiles of purified cells were obtained by RNA-seq, and cells intended for ChIP-seq were crosslinked upon single-cell suspension, prior to sorting, to minimize the impact on chromatin state. Cross-linked samples of each population were used for profiling histone modifications, such as H3K4me1 (candidate enhancers), H3K4me2 (promoters and enhancers), H3K4me3 (promoters), and H3K27ac (active enhancers). Finally, we identified open chromatin regions through ATAC-seq (Buenrostro et al., 2013; Lara-Astiaso et al., 2014).

Tissue-Resident Macrophages Can Be Distinguished by Their Gene Expression Patterns

To probe the spectrum of gene expression profiles among tissue-resident macrophages, we examined seven macrophage populations (brain microglia, spleen red pulp macrophages, liver Kupffer cells, lung macrophages, peritoneal cavity macrophages, and colonic large and ileal small intestinal macrophages), as well as monocytes. We identified 3,348 genes that were differentially expressed between at least two tissue-resident macrophages or monocytes (Figure 1B and Table S1). Many genes coding for TFs were uniquely expressed in specific macrophage populations, including *Sal1* in microglia and *Spi-c* in red pulp macrophages (Figures 1B and 1C). Other population-specific genes included *Clec4f* in Kupffer cells (Yang et al., 2013), *Car4* in lung macrophages, and *Tgfb2* in peritoneal macrophages (Figure 1C). Differentially expressed genes clustered into 11 groups by their expression patterns across samples (Figure 1B). Interestingly, macrophages that are presumably exposed to similar environmental cues displayed similar patterns of expression (Figures 1B, S1B, and S1C available online). Kupffer cells and splenic macrophages shared a cluster of highly expressed genes that are enriched for gene ontology (GO) annotations, such as heme and porphyrin metabolic processes, indicating their active role in erythrocyte turnover (Figures 1B and S1A, cluster III) (Chow et al., 2013). Similarly, small and large intestinal macrophages both express genes enriched for GO annotations that likely reflect their microbiota-exposed environment, such as response to bacterium and antigen processing (Figures 1B and S1A, cluster VII). RNA-seq analysis identified many genes expressed differentially by tissue-resident macrophage populations and monocytes and thereby corroborates and extends earlier expression profiling studies highlighting the inherent plasticity of this immune cell type (Gautier et al., 2012b).

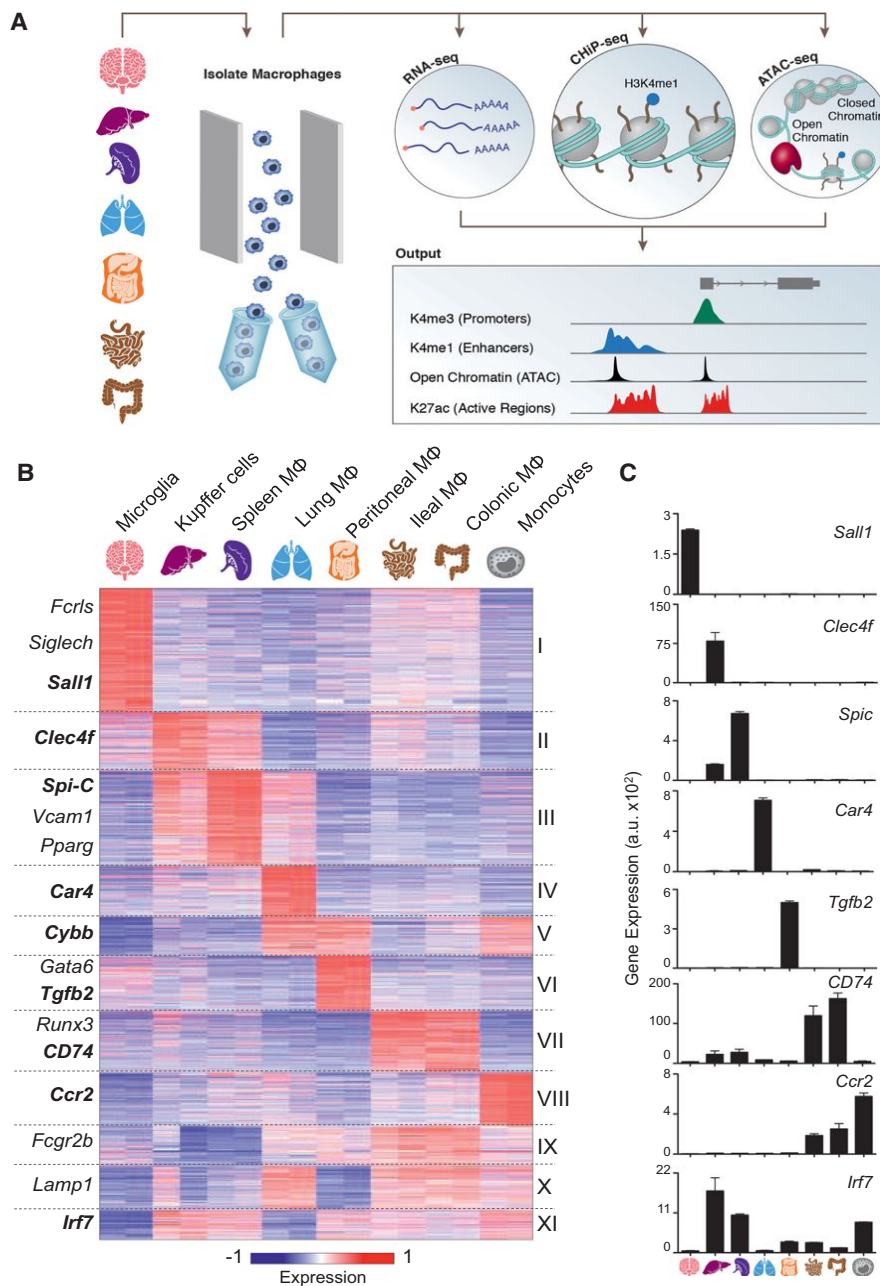


Figure 1. Tissue-Resident Macrophages Can Be Distinguished by Expression Patterns

(A) Schematic for defining the global regulatory elements of tissue-resident macrophages, monocytes, and neutrophils isolated by the gating strategy (Data S1) and analyzed by sequencing data from high-throughput RNA-seq, ChIP-seq, and ATAC-seq. A representative genome browser output is shown.

(B) K-means clustering ($K = 11$) of 3,348 differentially expressed genes in macrophages (Mφ) and monocytes.

(C) Bar graphs of individual gene expression in arbitrary units (a.u.). Error bars indicate SEM. See also Figure S1, Table S1, and Experimental Procedures.

Unique Regulatory Elements Distinguish Myeloid Cells

Profiling genome-wide histone modifications can shed light on the current regulatory state of a cell type. We focused on the distributions of H3K4me3, H3K4me1, and H3K27ac because, taken together, these marks classify three important functional elements: promoters, poised enhancers, and active enhancers. To better understand the mechanisms underlying macrophage plasticity, we compared the histone modification signal of circulating myeloid cells, specifically monocytes and neutrophils, to an average macrophage signature obtained by computationally merging all macrophage populations (Figure 2A). The promoter of *Mertk* is active in macrophages, but not monocytes or neutro-

phils as visualized by the “merged” macrophage H3K4me3 signal (Figure 2A). However, the majority of active promoters—regions with high H3K4me3 intensity—were shared by all myeloid cells (8,861 of 10,806 promoters, 82%; Figures 2B and S2B and Table S2A). On the other hand, of the total 30,976 putative enhancers, defined as regions distal to the transcriptional start site (TSS) with high H3K4me1 and low H3K4me3, only 8,260 (27%) were shared by macrophages, monocytes, and neutrophils. Indeed, both the loci of *Emr1* (F4/80) and *Mertk* feature intragenic enhancers unique to macrophages (Figures 2A and S2E). Active enhancers marked with H3K27ac were even less likely to be shared between all three myeloid cell types (10%; Figure S2A). For instance, the gene locus harboring the complement system genes (*C1qa*) contains a macrophage-specific H3K27ac-enriched region (Figures 2A and S2E). Macrophage-specific enhancers were less conserved than other myeloid enhancers, suggesting that they are a late-acquired evolutionary function (Figure S2D). Interestingly, of the 12,743 (7,825 active) macrophage-specific enhancers, rela-

tively few ($< 2\%$) are shared across all macrophage populations (Figure S2C). Because enhancer usage is highly differential across cell types when compared to promoters, the enhancer landscapes likely form the basis for macrophage specificity and plasticity.

To identify candidate regulators responsible for the distinct enhancer landscape of macrophages, monocytes, and neutrophils, we assessed these regions for enriched motifs and matched them to differentially expressed TFs (Table S2B). The TF-binding motif of PU.1 was common in all enhancer regions and was overrepresented in H3K4me1-marked regions shared by all ($p = 10^{-6}$) myeloid cell types (Figure 2C). Cell-type-specific

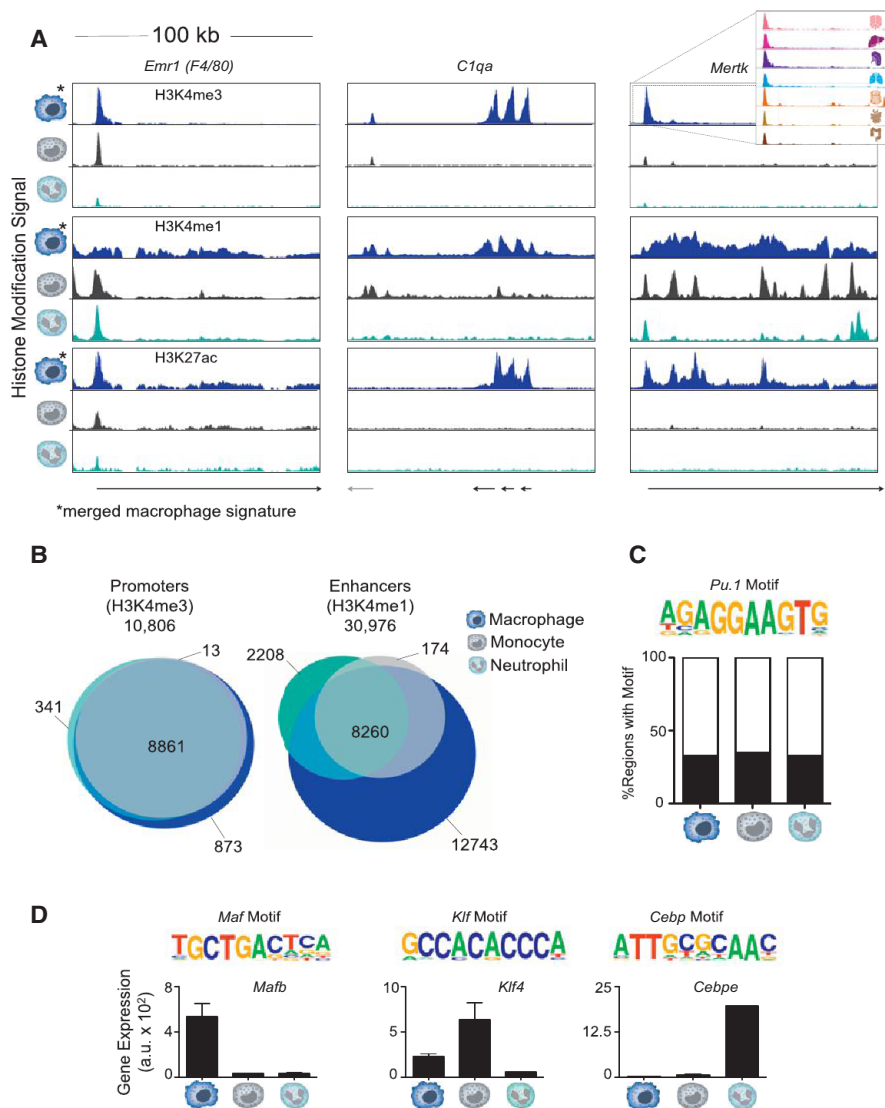


Figure 2. Comparing the Chromatin Landscape of Myeloid Cells Reveals Macrophage-Specific Enhancers

(A) Normalized profiles of H3K4me3, H3K4me1, and H3K27ac signal in 100 kilobase pair (kb) regions for monocytes, neutrophils, and “merged” macrophage signature of seven tissue-resident populations.

(B) Venn diagrams of the overlap of promoters (10,806; left) and enhancers (30,976; right) among macrophages, monocytes, and neutrophils.

(C) Representative motif of PU.1 enriched in H3K4me1-marked regions shared by all myeloid cell types ($p = 10^{-6}$) and percentage of regions with the motif in each cell type.

(D) Representative motifs of the indicated TF families found enriched in cell-type-specific H3K4me1-marked regions ($p \leq 10^{-5}$). Gene expression (a.u.) of the implicated family member. Error bars indicate SEM. See also Figure S2 and Table S2.

landscapes, despite their common lineage, indicating that the epigenome contributes significantly to cell-type specificity among myeloid cells.

Tissue-Resident Macrophages Exhibit Distinct Enhancer Landscapes that Reflect Ontogeny and Microenvironment

Because of the high divergence in tissue-resident macrophage expression and function, we next analyzed the chromatin landscape within the macrophage cell type. To determine the spectrum of enhancer usage across macrophage populations, we compared the chromatin profile of the 30,976 defined myeloid enhancers. We established that the vast majority of macrophage enhancers are

enhancers, on the other hand, tend to be bound by specific sets of TFs that regulate their chromatin states (Ghisletti et al., 2010; Heinz et al., 2010; Lara-Astiaso et al., 2014; Winter and Amit, 2014). Macrophage-specific enhancer regions exhibited significant overrepresentation of the Maf motif ($p = 10^{-23}$), in addition to high and specific expression of *Mafb* and *Maf* (Figures 2D and S2F), TFs known for their role in driving terminal macrophage differentiation (Aziz et al., 2009). Consistent with previous reports of *Klf4* as a monocyte regulator (Feinberg et al., 2007), enhancer regions unique to monocytes were enriched for the Klf motif ($p = 10^{-6}$), and *Klf4* was highly expressed in monocytes (Figures 2D and S2G). Finally, the Cebp motif was highly overrepresented in neutrophil-specific enhancers ($p = 10^{-30}$; Figure 2D). Of the Cebp family members implicated in myeloid cell differentiation, *Cebpe* is specifically required for neutrophil development (Yamanaka et al., 1997), and *Cebpe* is highly expressed in neutrophils (Figures 2D and S2H). Collectively, macrophages, monocytes, and neutrophils display highly distinct chromatin

unique to a small subset of populations (Figures 3 and 4). When we calculated the pairwise correlations of H3K4me1 read density between all samples, we found that the macrophage populations were more similar to each other than to neutrophils. However, the H3K4me1-marked enhancers in different macrophage populations display a high level of variation, especially when compared with H3K4me3 promoter signal (Figure 4A). The reproducibility and tissue specificity of macrophage enhancers were verified by biological replicates (mean $r = 0.885$). Moreover, another mark of enhancer usage, H3K4me2, uniquely predicted H3K4me1-marked regions within a sample with high accuracy (Figure S3D). This variation is representative of the distinct set of enhancers utilized by individual macrophage populations (Figures 4A and 4B).

Indeed, when the catalog of enhancers was clustered by H3K4me1 intensity across the populations, nearly all tissue-resident macrophages demonstrated uniquely utilized enhancers (Figure 4B, clusters I–III, V, VIII, and X). For instance, the

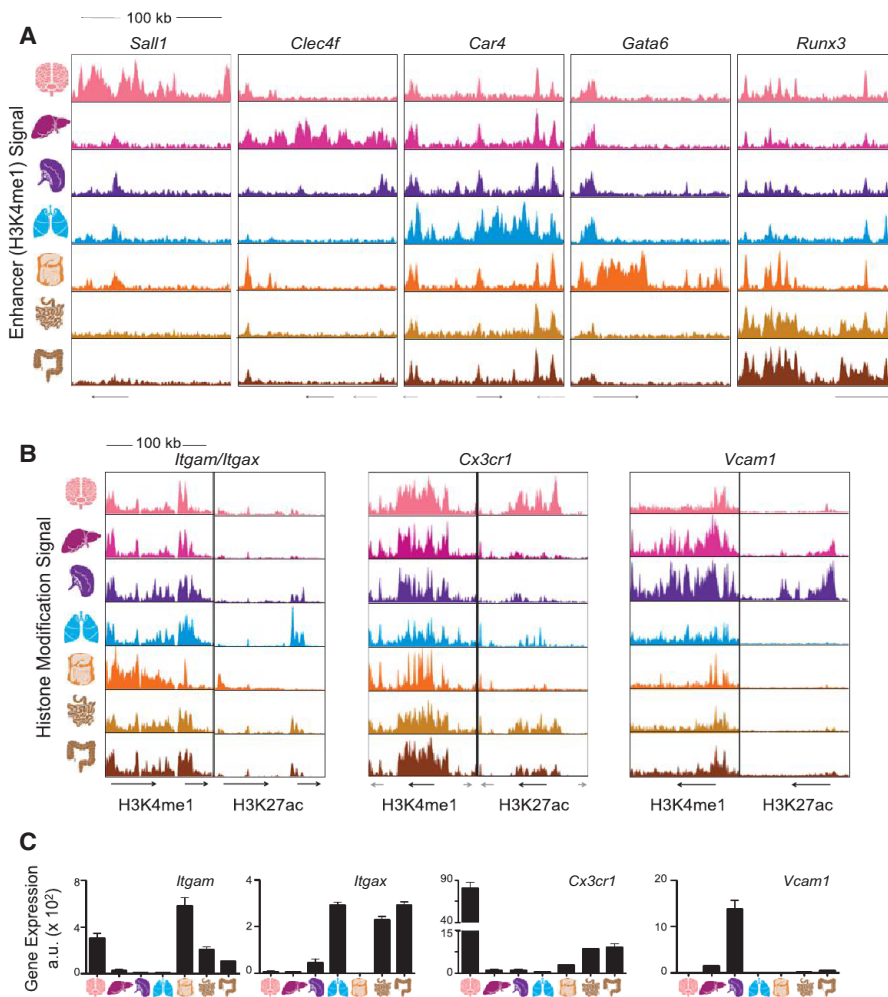


Figure 3. Tissue-Resident Macrophage Populations Have Unique Poised and Active Enhancers

(A) Normalized profiles of H3K4me1 signal for seven tissue-resident macrophage populations in 100 kb regions containing tissue-specific enhancers around the indicated genes.

(B) Normalized profiles of H3K27ac (right) signal in H3K4me1-marked (left) 100 kb regions containing the indicated genes.

(C) Bar graphs of expression (a.u.) for genes in loci of (B). Error bars indicate SEM. See also Figure S4.

opposed to the other tissue-resident macrophages analyzed, are mostly monocyte derived (Bain et al., 2014; Bogunovic et al., 2009; Varol et al., 2009). Accordingly, many more monocyte enhancers remain open in intestinal macrophages (85% and 87%) compared to other tissue-resident macrophages (chi-square $p < 10^{-5}$; Figure 4D). Conversely, embryo-derived macrophages shared many enhancers that are not present in intestinal macrophages or monocytes, including enhancers close to *Rxra* and *Marco* (Figure 4B, cluster XVII).

Analysis of the enhancer landscape also exposes relationships among tissue-resident macrophages that are unrelated to ontogeny. For instance, Kupffer cells and spleen macrophages cluster together as a result of the many enhancers they share (cluster IV, Figure 4C), likely reflecting the impact of their similar

environments characterized by prominent erythrocyte exposure. In addition, microglia and lung macrophages are most distant from other macrophages in the hierarchical tree and are excluded from an enhancer cluster shared by all other populations (cluster XV, Figure 4C).

Collectively, the distinct enhancer usage of different tissue-resident macrophages highlights how chromatin state dynamics allow for plasticity within a given cell type. Although some patterns reflect development, the level of variability between macrophage populations extends beyond this dimension. Thus, ontogeny, together with local microenvironment, likely contributes to shaping the enhancer landscape of tissue-resident macrophages.

intergenic region of *Sall1* exhibited a microglia-specific region enriched for H3K4me1 (cluster I; Figure 3A). Enhancers unique to lung macrophages were enriched for genes involved in metabolism of lipids and lipoproteins, indicating the role of these cells in surfactant lipoprotein metabolism (Hussell and Bell, 2014) (Table S3, cluster VIII). Moreover, an H3K4me1-marked region exclusive to peritoneal macrophages was detected within the *Gata6* gene locus (cluster X; Figure 3A). The highly similar small and large intestinal macrophages shared enhancers utilized in no other populations, including the regions adjacent to *Runx3* (cluster XI; Figure 3A). In confirmation, these clusters were reproduced by H3K4me2 intensity in the same regions, which was highly correlated with H3K4me1 across all analyzed macrophage populations (Figures S3A and S3B). In general, unique macrophage clusters are associated with tissue-specific genes and related functions (Table S3), indicating a potential role for the microenvironment in shaping chromatin dynamics.

Hierarchical clustering based on H3K4me1-marked or H3K4me2-marked enhancers positioned the monocytes next to the junction of small and large intestinal macrophages (Figures 4C and S3C). This suggests that ontogenic relationships influence chromatin dynamics because intestinal macrophages, as

environments characterized by prominent erythrocyte exposure. In addition, microglia and lung macrophages are most distant from other macrophages in the hierarchical tree and are excluded from an enhancer cluster shared by all other populations (cluster XV, Figure 4C).

Collectively, the distinct enhancer usage of different tissue-resident macrophages highlights how chromatin state dynamics allow for plasticity within a given cell type. Although some patterns reflect development, the level of variability between macrophage populations extends beyond this dimension. Thus, ontogeny, together with local microenvironment, likely contributes to shaping the enhancer landscape of tissue-resident macrophages.

Poised Enhancers Reflect Both Developmental Origin and Tissue Specificity

H3K4me1-marked enhancers are considered either active or poised, as demarcated by the presence or absence of the H3K27ac mark (Creighton et al., 2010). To further explore enhancer activity, we analyzed the distribution of H3K27ac in H3K4me1-marked enhancers (Figure S4). For example, the enhancer regions in the *Itgax* locus encoding the integrin CD11c are marked with H3K4me1 in all macrophages, although

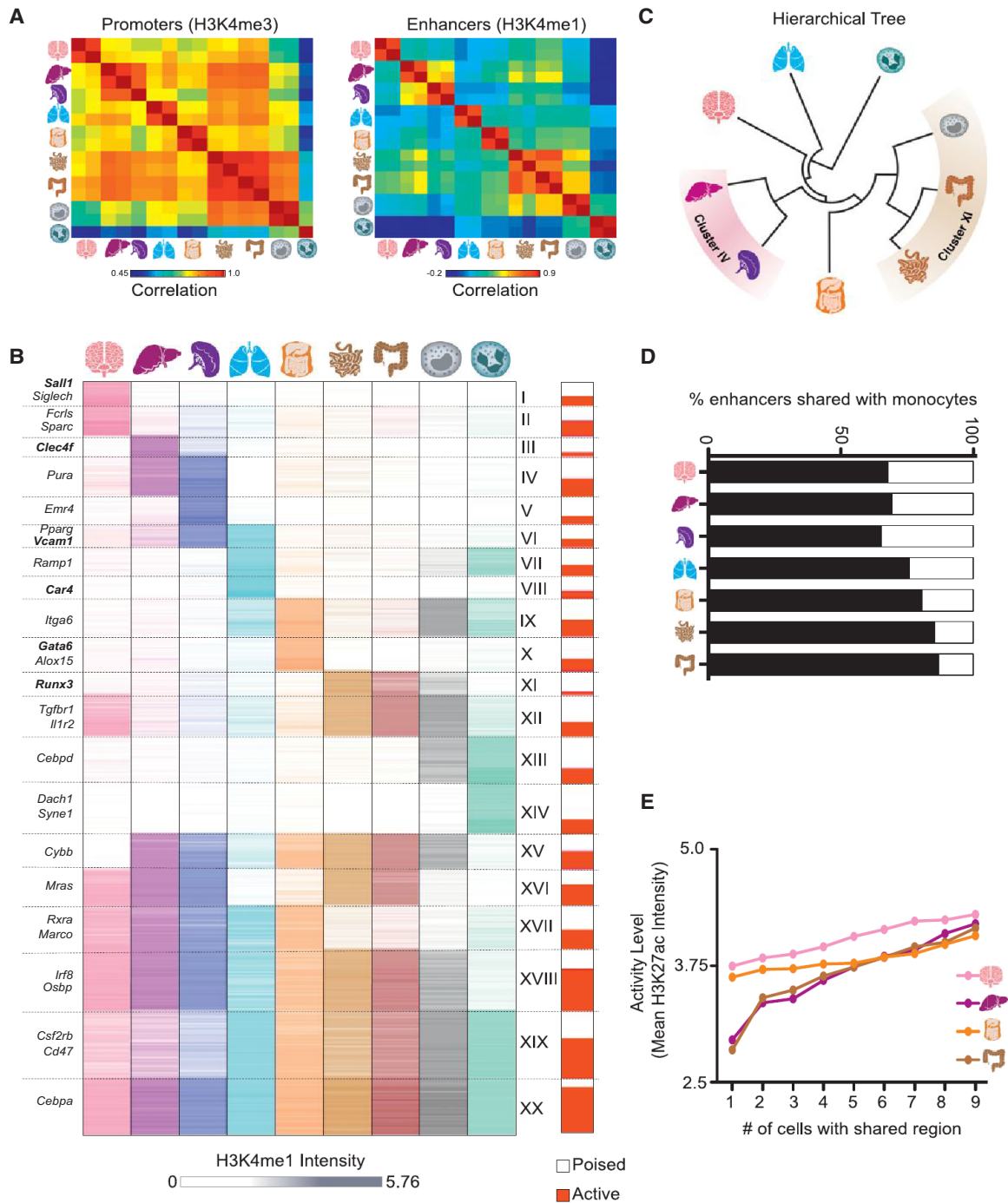


Figure 4. Tissue-Resident Macrophages Have Distinct Sets of Enhancers Determined by Ontogeny and Microenvironment

(A) Pairwise correlations between replicates with respect to H3K4me3 read density in promoters (left) and H3K4me1 read density in enhancers (right).

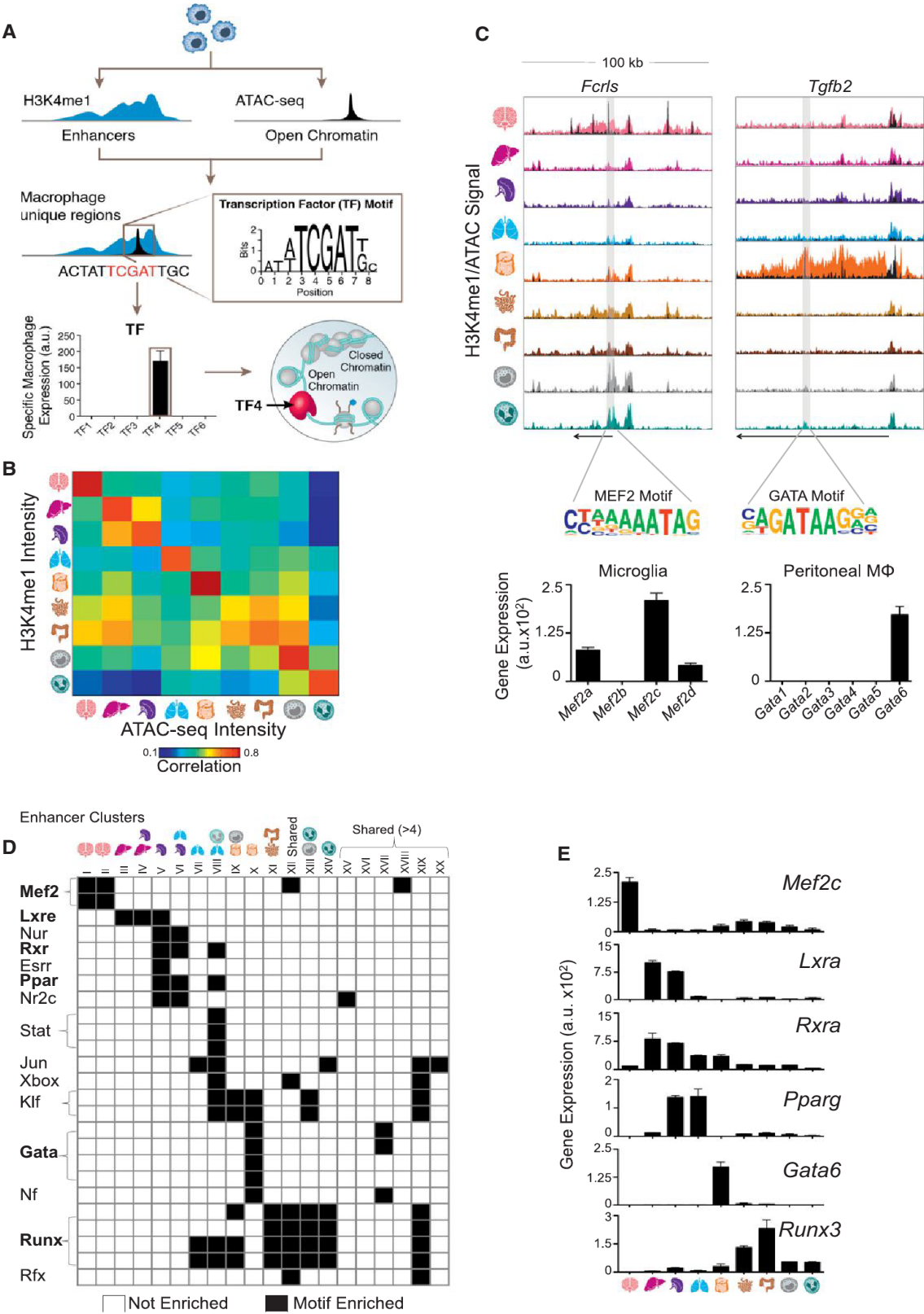
(B) K-means clustering (K = 20) of the H3K4me1 intensity in 30,976 high-confidence enhancer regions. The proportion of enhancers active (H3K27ac high) in at least one sample in each cluster is shown in red (right bar).

(C) Hierarchical tree resulting from clustering on H3K4me1 intensity in enhancer regions.

(D) Percentage of H3K4me1-marked regions in monocytes shared with each tissue-resident macrophage.

(E) Line plots showing that enhancer activity level (mean H3K27ac intensity) increases with the number of samples that share the H3K4me1-marked regions.

See also Figures S3 and S4 and Table S3.



(legend on next page)

expression of *Itgax* is restricted to lung and intestinal macrophages. This is reflected in the H3K27ac signal denoting transcriptional activity in only these macrophage populations (Figures 3B and 3C). Likewise, *Cx3cr1*, a gene encoding a chemokine receptor, exhibits a H3K4me1-marked region shared by all macrophages but is activated only in intestinal macrophages and microglia (Figures 3B and 3C). On the other hand, the region surrounding the *Vcam1* gene, along with the previously described loci (Figure S4A), exhibits tissue-specific macrophage enhancers with H3K4me1 signal in proportion to its H3K27ac activity and expression (Figures 3B and 3C). Thus, analysis of both H3K4me1 and H3K27ac in enhancers allows us to describe the current activity, developmental origin, and potential of these cells to activate various gene sets under external stimuli.

We determine enhancer activity based on the level of H3K27ac intensity in H3K4me1-marked candidate enhancers and classify poised enhancers as those that have low H3K27ac in all populations (Experimental Procedures). On a global scale, H3K27ac intensity correlates with H3K4me1 in enhancers and exhibits similar relationships between populations (Figures S4B–S4E). However, individual enhancers vary in the relationship between the two marks as evidenced by differing proportions of active and poised enhancers in the H3K4me1 clusters (Figure 4B, right bar). Dissecting this pattern further, we find that enhancers shared by multiple macrophage populations have a higher level of activity than those that are shared by few or none (Figure 4E). For instance, enhancers that were present in all but one of the populations were significantly more likely to be active in Kupffer cells than those enhancers shared with only one other population (kstest $p = 9.35 \times 10^{-37}$; Figure 4E). This trend held true for all samples except neutrophils (Figure S4F). Therefore, plasticity of tissue-resident macrophages may arise from selective activation of developmentally derived enhancers or de novo enhancers triggered by the local microenvironment.

Tissue Regulators Define Distinct Sets of Macrophage Enhancers

We hypothesized that the distinct enhancer landscape in tissue-resident macrophages is the result of the combinatorial action of tissue regulators. Assuming that the expression of TFs should match the utilization of the enhancers that they regulate, we established a computational pipeline to identify candidate regulators for each enhancer cluster (Figure 5A). Because H3K4me1-marked enhancers may span several kilobases, but TF motifs usually occupy no more than a dozen bases, we generated ATAC-seq peaks corresponding to the same tissue-resident macrophages to narrow our search regions to the likely site of

TF binding (Buenrostro et al., 2013; Lara-Astiaso et al., 2014). Marking open chromatin, ATAC-seq intensity is highly correlated with H3K4me1 intensity in enhancers, and the vast majority of putative enhancer regions contain at least one ATAC-seq peak (Figures 5B, S5C, and S5D). For each cluster, we overlapped the enhancer coordinates with the ATAC-seq peak and lifted the DNA sequence from these regions to search for significantly enriched, cluster-specific motifs. Finally, to determine the most likely regulator from the TF family that matched the motif, we compared the gene expression profile of all family members to the enhancer signature. In this manner, we identified candidate regulators for many of the clusters shown in Figure 4B (Figures 5C–5E, S5E, and S5F and Table S4).

For example, we found that the GATA motif was overrepresented in the cluster of enhancers specific to peritoneal macrophages (cluster X). Among the GATA TF family members, *Gata6* was the most highly and differentially expressed in peritoneal macrophages (Figures 5C–5E). Therefore, *Gata6* is a likely regulator of peritoneal-specific macrophage enhancers, as supported by recently published results (Gautier et al., 2014; Okabe and Medzhitov, 2014; Rosas et al., 2014). Indeed, GATA motifs appear within enhancers associated with genes expressed exclusively in peritoneal macrophages, such as *Tgfb2* (Figures 1A and 5C). Likewise, MEF2 binding motifs were overrepresented in microglia-specific enhancer clusters and appeared at enhancers of microglia-specific genes such as *Fcrls* (Figures 5C, 5D, and 4B, clusters I and II). *Mef2c* was implicated as the mostly highly and differentially expressed MEF2 family member in microglia (Figures 5C–5E and Table S4). Other candidate TFs include *Lxra* in Kupffer cells and spleen macrophages (clusters III and V) and *Pparg* in spleen and lung macrophages (clusters V and VI). Intestinal macrophages, along with monocytes and neutrophils, are enriched for RUNX family motifs, with *Runx3* highly expressed in intestinal macrophages (clusters XI–XIV) (Figures 5D and 5E and Table S4).

Our data suggest that the distinct enhancer landscapes of tissue-resident macrophages result from the restricted expression and binding of TFs. The orchestration of chromatin modifications is regulated by crosstalk between the environment and ontogeny through a small number of TFs (Heinz et al., 2010). Hence, macrophage identity is shaped by tissue-restricted TFs in conjunction with those ubiquitously present in macrophages, such as PU.1.

Macrophage Enhancer Landscapes Are Imparted by the Microenvironment

To assess the contribution of environmental signals as opposed to developmental processes, we assessed the chromatin state

Figure 5. Identification of Candidate Regulators in Tissue-Resident Macrophage Enhancers Using ATAC-Seq

(A) Schematic of pipeline to identify candidate regulators of enhancers. For each cluster, enhancers are matched to ATAC-seq peaks, genomic sequence is lifted for input into the motif finder, and the enriched motif is matched to the TF family member with corresponding expression.

(B) Pairwise correlations between H3K4me1 and ATAC-seq intensity in enhancer regions.

(C) Normalized profiles of H3K4me1 signal in 100 kb regions with ATAC-seq peaks overlaid in black, containing tissue-specific enhancers around the indicated genes. Shaded regions indicate location of the relevant motif. Bar graphs of gene expression (a.u.) for TF family members of *Mef2* in microglia and *Gata* in peritoneal macrophages. Error bars indicate SEM.

(D) Heatmap of significantly enriched motifs ($p \leq 10^{-5}$) in H3K4me1-marked regions from each cluster in Figure 4B (see Table S4).

(E) Bar graphs show gene expression (a.u.) for candidate TFs that match motifs in (D). Error bars indicate SEM.

See also Figure S5 and Table S4.

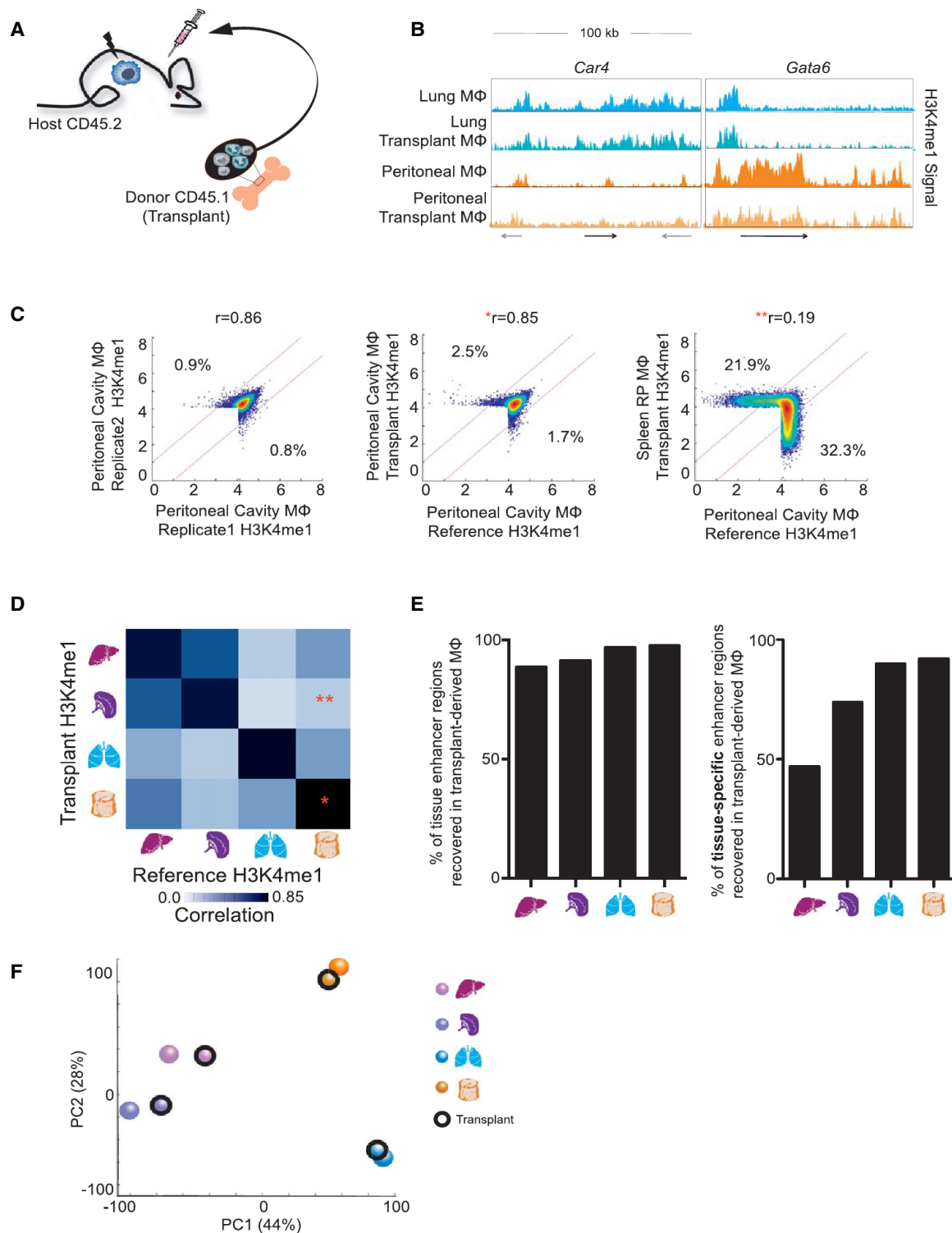


Figure 6. Microenvironment Signals Shape the Enhancer Landscape of Adult Bone-Marrow-Derived Macrophages

(A) Schematic of bone marrow transplant experiment.

(B) Normalized profiles of H3K4me1 signal of reference macrophages and transplant-derived macrophages both showing tissue-specific enhancers around *Car4* and *Gata6* in 100 kb regions.

(legend continued on next page)

of macrophages derived from transplanted adult bone marrow (BM) that replace endogenous embryo-derived tissue-resident macrophages upon lethal irradiation (Ginhoux et al., 2010; Hashimoto et al., 2013; Schulz et al., 2012; Virolainen, 1968; Yona et al., 2013) (Figure 6A). Four months after engraftment, we retrieved the donor transplant-derived lung, spleen, liver, and peritoneal macrophages using the congenic CD45 alleles to distinguish them from host macrophages (Figure S6A; Experimental Procedures). Critically, we found that macrophages from transplanted adult BM acquire enhancers found in embryonic macrophages in a tissue-specific manner, including those in the *Car4* locus of lung macrophages and the *Gata6* locus found in peritoneal macrophages (Figure 6B).

Importantly, the H3K4me1 intensity of adult transplant-derived macrophages was highly reproducible and much more similar to their embryo-derived counterpart than any other macrophage population (Figures 6C, 6D, and S6B–S6D). Macrophages derived from donor BM recovered 89%–98% overall and 47%–92% of tissue-specific enhancers from the reference (Figure 6E, Experimental Procedures). Similarly, principal component analysis (PCA) revealed that transplanted macrophages were positioned closer to their corresponding reference macrophage than other reference or transplanted macrophages (Figure 6F). Similar results were seen with active enhancer (H3K27ac) regions (Figures S6E–S6H). Collectively, these data establish that adult bone marrow precursors can acquire the tissue-specific enhancer landscape of macrophages seeded in the tissues during development, reinforcing that the tissue microenvironment plays a prominent and active role in establishing macrophage identity.

Differentiated Tissue-Resident Macrophages Can Be Reprogrammed by the Tissue Microenvironment

Typically, once the chromatin landscape is specified during differentiation, cells lose the plasticity to revert to their original precursor state or convert—without artificial manipulation—into other cell types. Based on the results of the above transplant experiment, we asked whether differentiated tissue-resident macrophages would retain sufficient plasticity to adapt to a new microenvironment upon transfer into an ectopic tissue.

Lung and peritoneal macrophages display distinct expression profiles (Figure 7A). Moreover, they harbor distinct chromatin landscapes with tissue-specific enhancers and candidate regulators, such as *Gata6* and *Pparg* (Figures 7B, 7C, 5E, and S7A). To test the potential of the environment to reprogram differentiated macrophages, we sorted peritoneal macrophages from CD45.1⁺ donor mice and transferred them intratracheally into the alveolar cavity of CD45.2⁺ animals (Figures 7D and S7B).

Although the engrafted macrophages retained high levels of CD11b surface expression, akin to endogenous peritoneal macrophages, ICAM2—a peritoneal macrophage-specific surface marker (Gautier et al., 2012b) (Figure 7A)—was downregulated upon transfer to levels on par with lung macrophages (Figure 7E). Transferred macrophages upregulated lung macrophage-specific genes, including *Chi3l3*, *Sftpc*, *Car4*, and the TF *Pparg* (Figure 7F), while downregulating peritoneal macrophage-specific genes, including *Alox15* and the TF *Gata6* (Figure 7F). Global RNA-seq analysis of transferred cells demonstrated that 70% of genes highly and differentially expressed in peritoneal or lung macrophages (708 of 1,014) had switched from a peritoneal cavity profile to resemble lung macrophages (Figure 7G and Table S5), resulting in a significant shift in the distribution of genes ($p = 5.8 \times 10^{-37}$; $p = 2.9 \times 10^{-37}$; Figure S7C). The overall transition was confirmed by PCA analysis of the expression profiles, which placed transferred cells closer to lung than peritoneal macrophages (Figure 7H). This transfer experiment indicates that differentiated tissue-resident macrophages retain their plasticity and further emphasizes the critical role of the microenvironment in shaping the functional identity of steady-state macrophages.

DISCUSSION

The epigenomic state of a cell regulates gene expression, differentiation, and cellular identity. We show that, at least in the case of tissue-resident macrophages, the chromatin landscape provides a mechanism for plasticity and allows for crosstalk between the environment and tissue-specific macrophage function. We provide a comprehensive resource of the annotated regulatory elements, including promoters (H3K4me3), poised enhancers (H3K4me1), and active enhancers (H3K27ac) of seven different tissue-resident macrophage populations, as well as monocytes and neutrophils. We demonstrate that tissue-resident macrophage populations are distinct with respect to their set of regulatory elements. Much of the previous work on macrophage chromatin states explored either BM culture-derived macrophages or thioglycollate-elicited monocyte-derived peritoneal macrophages. Although these studies offered many insights on general macrophage function, our results show that the application of these earlier findings to all macrophage populations cannot be taken for granted. Rather, tissue-resident macrophages are subject to an additional level of regulation by a combination of ontogeny and the dynamic influence of the tissue microenvironment, which together shape their distinct chromatin landscapes.

Tissue-resident macrophages have distinct global expression profiles, as we show using RNA-seq in confirmation of prior

(C) Density scatterplots of H3K4me1 intensity in H3K4me1-marked regions for reference peritoneal macrophages (x axis) between replicates (left), compared to peritoneal transplant-derived macrophages (middle), or spleen transplant-derived macrophages (right). Regions of 2-fold differential (pink lines) and Pearson's correlation between samples are indicated.

(D) Pairwise correlations between transplant and reference macrophages with respect to H3K4me1 intensity in original and novel enhancers. The single asterisk (*) and double asterisk (**) correspond to the data shown in the respective plots from (C).

(E) Percent of total (left) or tissue-specific (right) reference macrophage H3K4me1-marked regions recovered by transplant-derived macrophages.

(F) PCA of H3K4me1 intensity showing transplanted macrophages group with their respective reference macrophages.

See also Figure S6.

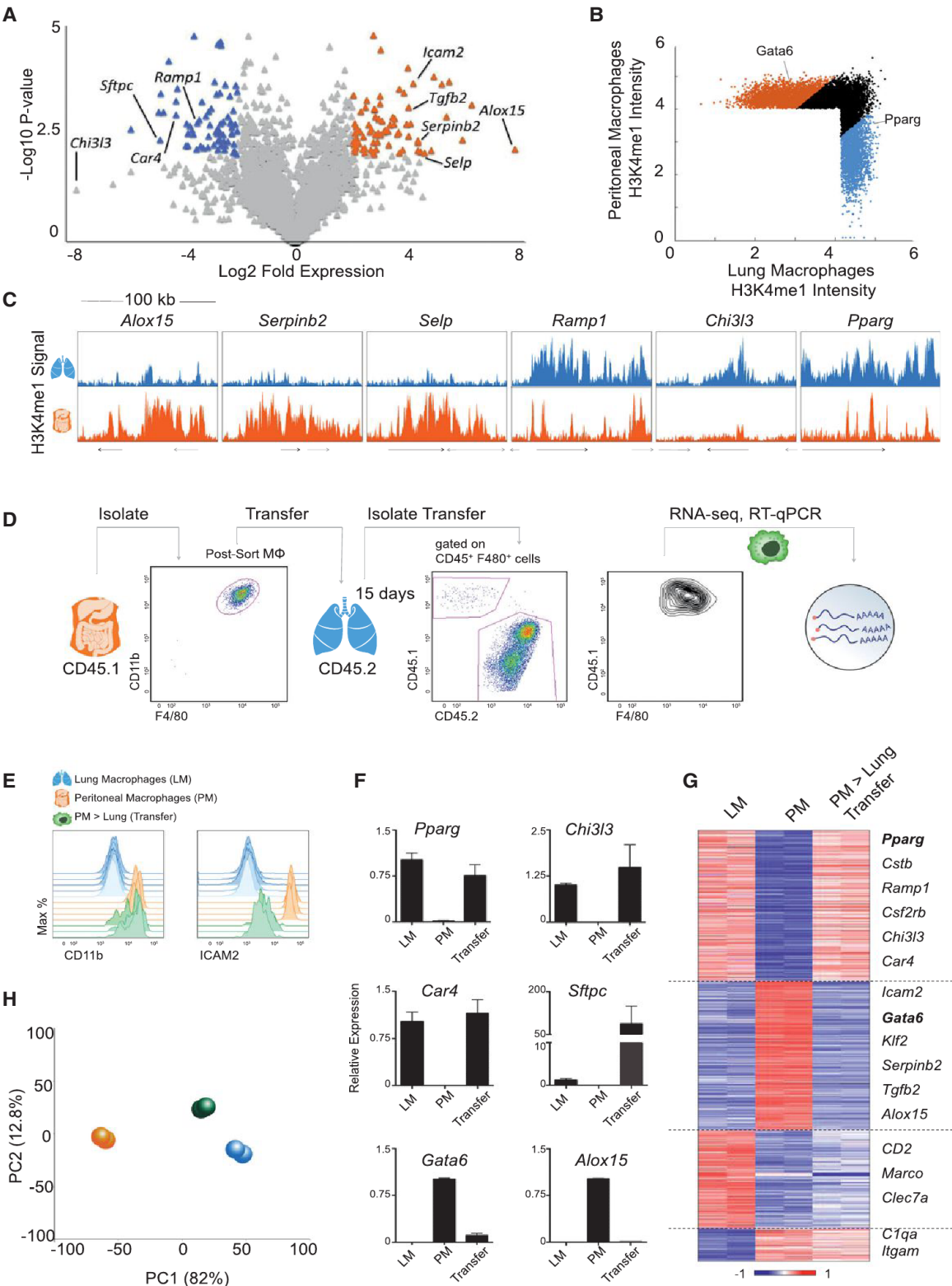


Figure 7. Differentiated Tissue-Resident Macrophages Are Reprogrammed by a New Microenvironment
(A) Volcano plot of relative expression in peritoneal (PM; right) and lung (LM; left) macrophages with differential (> 4-fold), statistically significant ($p < 0.01$) genes indicated in orange and blue, respectively.

(legend continued on next page)

results obtained on a smaller scale (Gautier et al., 2012b). To assess the extent of artificial induction of stimulus-response genes in our RNA-seq data, we compared the expression profiles of the tissue-resident macrophage populations with stimulated dendritic cells across multiple time points (Garber et al., 2012). Because the highest correlation was with unstimulated cells, we conclude that our protocols did not cause undue activation of immediate early genes (Figure S2D). Chromatin states, moreover, are stable within the time span of macrophage isolation (Experimental Procedures) (Garber et al., 2012; Ostuni et al., 2013); thus, the in vivo chromatin states are even less likely to be disrupted.

Through assessing enhancer activity, we add a new dimension to understanding how the identities of tissue-resident macrophages are regulated. Some poised enhancers are inherited in development and thus encode ontogenic memory (Nord et al., 2013); such enhancers are ubiquitously marked with H3K4me1 but may be active only in a subset of populations (Stergachis et al., 2013). Lineage-tracing experiments show that *Cx3cr1*, which displays a poised enhancer already at the common myeloid precursor (CMP) stage (Lara-Astiaso et al., 2014), is expressed during development of all macrophage populations (Yona et al., 2013). Interestingly, the *Cx3cr1* locus retains H3K4me1-marked enhancers in all tissue-resident macrophages populations, although it is almost exclusively acetylated and expressed in microglia and intestinal macrophages. De novo enhancers encompass much of the diversity between different tissue-resident macrophage populations and provide a possible account for macrophage plasticity. Many of these regions are poised and, along with latent enhancers not preemptively marked for activation, reflect the potential to respond to future stimuli (Ostuni et al., 2013). Indeed, we show that the highest percentages of poised enhancers are at regions present in only one or two macrophage populations, demonstrating complex regulatory mechanisms, which can only be analyzed comprehensively on the level of the chromatin.

Master regulators, such as PU.1, work in combination with other TFs to specify active regulatory regions in a cell-type-specific manner (Heinz et al., 2010). Here, we implicate Maf family TFs, such as Maf and MafB, in the establishment of macrophage-specific regulatory regions. Within macrophages, we further identified candidate regulators that may work in combination to establish the distinct chromatin landscape for each tissue-resident macrophage population. These TFs, which include Mef2c in microglia (Speliotis et al., 1996), Lxra in Kupffer cells and splenic macrophages (A-Gonzalez et al., 2013; Joseph et al., 2004), Pparg in splenic red pulp and lung macrophages (Gautier et al., 2012a; Schneider et al., 2014), Gata6 in peritoneal macrophages (Okabe and Medzhitov, 2014), and Runx3 in intestinal macrophages, have varying support in the literature.

Future studies should focus on knocking out or overexpressing these candidate regulators and observing the effects on the regulatory networks, epigenomic landscape, and identity of specific macrophage populations.

These candidate regulators are likely to be the conduit through which signals in the tissue microenvironment influence the chromatin landscape of macrophages. Our results are consistent with previous studies showing that heme and retinoic acid promote the induction of the tissue-specific macrophage TFs, Spi-C and Gata6, respectively (Haldar et al., 2014; Okabe and Medzhitov, 2014). Notably, some of the TFs we identified, such as Klf4 and Pparg, have been reported to be associated with functional macrophage polarization (Murray et al., 2014). However, given the heterogeneity of enhancer landscapes and the tissue-specific signatures of poised enhancers, the activation of tissue-resident macrophages is likely to be more complex than previously anticipated. We confirmed the role of the environment in shaping macrophage identity by enforcing replacement of endogenous, embryo-derived macrophages with adult BM cells, and almost all enhancers were recovered in a tissue-specific manner. The question of whether monocytes contribute to the embryo-derived macrophage population in the case of nonirradiating injury or infection in the long-term remains unresolved (Epelman et al., 2014; Ginhoux and Jung, 2014). However, our study indicates that, if precursors do enter and remain in the tissue, they will be able to assimilate to the local macrophage population. Moreover, they adopt both poised and active regulatory elements, indicating their functionality.

Interestingly, when we transferred fully differentiated macrophages to an alternate tissue, we found that the new environment was sufficient to reshape their expression. The chromatin landscape is, therefore, specialized within the tissue, while still retaining the capacity to be reversed (as seen in Okabe and Medzhitov, 2014) and reprogrammed. This also presents the possibility that peritoneal macrophages, or other differentiated macrophages, could potentially serve as a therapeutic source of macrophages, as recently shown for precursors (Happle et al., 2014; Suzuki et al., 2014).

Macrophages are important sentinels of the immune system, embedded in each tissue, and are responsive to local changes in the microenvironment. Identifying the regulatory elements that define each tissue-resident macrophage identity is, therefore, critical for understanding their role in alerting the immune system to disease. The means by which the dysregulation of chromatin landscape and transcription factors can lead to impaired function directs the exploration of new therapeutic strategies to stimulate or inhibit the appropriate, tissue-specific macrophage response. Finally, a similar study in humans will be highly beneficial for uncovering the mechanisms of regulatory aberrations in immunological disorders.

(B) Scatterplot comparing intensity in H3K4me1-marked regions with differential PM (orange) and LM (blue) enhancers indicated.

(C) Normalized profiles of H3K4me1 signal of PM and LM in 100 kb regions.

(D) Schematic of transplant experiment: representative FACS plot of purified CD45.1⁺ PM prior to transfer and retrieved from host CD45.2⁺ lung.

(E–H) Resident LM, PM, and transferred macrophages recovered from the host lung tissue (PM > lung transfer) were analyzed by flow cytometry (E) and qRT-PCR

(F). Error bars indicate SEM. RNA-seq from transferred macrophages was compared to reference macrophages for 1,014 differential genes sorted in a heatmap (G) and analyzed by PCA (H).

See also Figure S7 and Table S5.

EXPERIMENTAL PROCEDURES

Mice

C57BL/6 mice were purchased from Harlan (Rehovot, Israel). Macrophages were isolated from 6- to 7-week-old females. All animals were kept in specific pathogen-free (SPF) conditions and handled according to the protocols approved by the Weizmann Institute Animal Care Committee as per international guidelines.

Tissue Processing, Flow Cytometry, and Cell Sorting

Macrophages were purified following protocols (Extended Experimental Procedures; Data S1) as previously described (Gautier et al., 2012b; Zigmond et al., 2012). Cells were isolated directly from mouse tissue, stained, and sorted on an Aria III either into lysis/binding buffer (Life Technologies) for RNA-sequencing or crosslinked in 1% formaldehyde (Pierce Biotechnologies) prior to sorting into PBS for ChIP-seq.

Bone Marrow Transplant and Intratracheal Transfer

Cells were isolated from CD45.1 mice by bone marrow flushing and injected IV into irradiated 6-week-old CD45.2 females. Macrophages were isolated from tissues and processed as above 11–12 weeks later. For intratracheal transfer, peritoneal cavity macrophages were sorted from CD45.1 animals. 5×10^5 cells in sterile PBS or PBS alone were placed in the distal oral cavity of anesthetized 5-week-old animals. Cells were isolated after 15 days from the lung and processed as above.

Quantitative PCR

mRNA was isolated with Dynabeads oligo(dT) (Life technologies), and reverse-transcribed to cDNA using AffinityScript RT kit (Agilent). Quantitative PCR (qPCR) was performed with LightCycler480 SYBR Green I Master Mix (Roche) in triplicate, normalizing to Actb. Primers are listed in Table S6.

RNA Isolation, Library Construction, and Analysis

10^4 – 10^5 cells from each population were sorted into 100–200 μ l of lysis/binding buffer (Life Technologies). mRNA was captured with 12 μ l of Dynabeads oligo(dT) (Life Technologies), washed, and eluted at 70°C with 10 μ l of 10 mM Tris-Cl (pH 7.5). We used a derivation of MARS-seq as described (Jaitin et al., 2014), developed for single-cell RNA-seq to produce expression libraries with a minimum of two replicates per population. Table S7 shows MARS-seq primers. We sequenced an average of 4 million reads per library and aligned them to the mouse reference genome (NCBI 37, mm9) using TopHat v2.0.10 (Trapnell et al., 2009) with default parameters. Expression levels were calculated and normalized using ESAT software (<http://garberlab.umassmed.edu/software/esat>). RNA-seq analysis in Figure 1 focused on genes in 25th percentile of expression with a 2-fold differential between at least two populations. For transferred macrophages in Figures 7G and 7H, highly expressed and differential (2-fold) peritoneum and lung macrophage genes were analyzed. Details are provided in the Extended Experimental Procedures.

High-Throughput-ChIP-Seq

10^5 crosslinked cells were used for ChIP-seq, as described (Blecher-Gonen et al., 2013; Garber et al., 2012). Following crosslinking and sorting, chromatin was fragmented by sonication, and the mixture was purified with magnetic beads (Invitrogen, Dynabeads) conjugated to 1 ng of H3K4Me1 (Abcam and Millipore), H3K4Me2 (Abcam), H3K4Me3 (Millipore), or H3K27Ac (Abcam) antibodies. After barcoding, pooled DNA was sequenced (HiSeq 1500, Illumina) to achieve a minimum of 10^7 aligned reads per sample.

ATAC-Seq

To profile open chromatin, we used the ATAC-seq protocol developed by Buenostro et al. (2013) with modifications described in the Extended Experimental Procedures and by Lara-Astiaso et al. (2014).

Processing of ChIP-Seq, ATAC-Seq, and Chromatin Peak Calling

Reads were aligned to the mouse reference genome (mm9, NCBI 37) using Bowtie aligner version 1.0.0 (Langmead et al., 2009) with best match parameters (bowtie -m 1 -sam-best-strata -v 2). To identify regions of enrichment,

peaks, from ChIP-seq reads of H3K4me1 and H3K4me3, we used the HOMER package *makeTagDirectory* followed by *findPeaks* command with the histone parameter (Heinz et al., 2010). Union peaks file were generated by combining and merging overlapping peaks in all samples.

Chromatin Analysis

The read density (number of reads in 10 million total reads per 1,000 bp) was calculated in each region from union peaks files for H3K4me1 and H3K4me3. We quantile normalized the average read density of replicates in high-confidence regions (i.e., above threshold in both replicates; Extended Experimental Procedures and Data S1) of each cell type. The region intensity was given in log-base2 of the normalized density ($\log_2(x+1)$). We define promoters as these 10,806 H3K4me3 regions and enhancers as the 30,976 H3K4me1 nonpromoter regions—i.e., neither overlapping H3K4me3 regions nor within $\pm 2,000$ bp of a TSS (Data S1). We calculated intensity of H3K27ac in each H3K4me1 region designated as an enhancer. To compare the relative activity of clusters, we classified 14,112 active enhancers as those that had density ≥ 15 before normalization (see below) in any cell type (Data S1). Details are provided in the Extended Procedures.

For motif finding, we independently called peaks in ATAC-seq, as above, and identified the maximum peak that overlapped each enhancer region. The overlapping sequences were input for HOMER package motif finder algorithm *findMotifGenome.pl* (Heinz et al., 2010). See also Extended Experimental Procedures.

Gene Tracks and Normalization

All gene tracks were visualized as *bigWig* files of the combined replicates normalized to 10,000,000 reads. The merged macrophage track was created by adding together the fastq reads of all macrophage populations and normalizing to total reads. For visualization, the tracks were smoothed by averaging over a sliding window of 500 bases.

ACCESSION NUMBERS

The main GEO accession number for the raw and processed sequencing data reported in this paper is GSE63341, with the subseries accession numbers GSE63338, GSE63339, and GSE63340 for ATAC-seq, ChIP-seq, and RNA-seq, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, one data file, and seven tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2014.11.018>.

ACKNOWLEDGMENTS

We thank members of the I.A., S.J., and M.M. labs for discussions. We thank Tali Wiesel and Genia Brodsky for artwork. Research in the I.A. lab is supported by the European Research Council (309788) and the Israeli Science Foundation (1782/11), The Human Frontiers Science Program, Career Development Award, and the Center for Excellence in Genome Science from the NHGRI 1P50HG006193. Research in the S.J. lab is supported by the Israeli Science Foundation (887/11), the European Research Council (340345), and the Deutsche Forschungsgemeinschaft (FOR1336). M.M. is supported by NIH R01CA154947A, NIHR01CA173861, NIHU01AI095611, and NIH-R01AI104848. D.W. is grateful to the Azrieli Foundation for the Azrieli Fellowship award and EMBO for the Long-Term Fellowship. Y.L. is grateful for the MSTP training grant support.

Received: October 15, 2014

Revised: November 11, 2014

Accepted: November 11, 2014

Published: December 4, 2014

REFERENCES

- A-Gonzalez, N., Guillen, J.A., Gallardo, G., Diaz, M., de la Rosa, J.V., Hernandez, I.H., Casanova-Acebes, M., Lopez, F., Tabraue, C., Beceiro, S., et al. (2013). The nuclear receptor LXR α controls the functional specialization of splenic macrophages. *Nat. Immunol.* **14**, 831–839.
- Abutbul, S., Shapiro, J., Szaingurten-Solodkin, I., Levy, N., Carmy, Y., Baron, R., Jung, S., and Monsonogo, A. (2012). TGF- β signaling through SMAD2/3 induces the quiescent microglial phenotype within the CNS environment. *Glia* **60**, 1160–1171.
- Aziz, A., Soucie, E., Sarrazin, S., and Sieweke, M.H. (2009). MafB/c-Maf deficiency enables self-renewal of differentiated functional macrophages. *Science* **326**, 867–871.
- Bain, C.C., Bravo-Blas, A., Scott, C.L., Gomez Perdiguero, E., Geissmann, F., Henri, S., Malissen, B., Osborne, L.C., Artis, D., and Mowat, A.M. (2014). Constant replenishment from circulating monocytes maintains the macrophage pool in the intestine of adult mice. *Nat. Immunol.* **15**, 929–937.
- Blecher-Gonen, R., Barnett-Itzhaki, Z., Jaitin, D., Amann-Zalcenstein, D., Lara-Astiaso, D., and Amit, I. (2013). High-throughput chromatin immunoprecipitation for genome-wide mapping of in vivo protein-DNA interactions and epigenomic states. *Nat. Protoc.* **8**, 539–554.
- Bogunovic, M., Ginhoux, F., Helft, J., Shang, L., Hashimoto, D., Greter, M., Liu, K., Jakubzick, C., Ingersoll, M.A., Leboeuf, M., et al. (2009). Origin of the lamina propria dendritic cell network. *Immunity* **31**, 513–525.
- Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218.
- Butovsky, O., Jedrychowski, M.P., Moore, C.S., Cialic, R., Lanser, A.J., Gabriely, G., Koeglsparger, T., Dake, B., Wu, P.M., Doykan, C.E., et al. (2014). Identification of a unique TGF- β -dependent molecular and functional signature in microglia. *Nat. Neurosci.* **17**, 131–143.
- Chow, A., Huggins, M., Ahmed, J., Hashimoto, D., Lucas, D., Kunisaki, Y., Pinho, S., Leboeuf, M., Noizat, C., van Rooijen, N., et al. (2013). CD169⁺ macrophages provide a niche promoting erythropoiesis under homeostasis and stress. *Nat. Med.* **19**, 429–436.
- Cirillo, L.A., Lin, F.R., Cuesta, I., Friedman, D., Jarnik, M., and Zaret, K.S. (2002). Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* **9**, 279–289.
- Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. USA* **107**, 21931–21936.
- Epelman, S., Lavine, K.J., and Randolph, G.J. (2014). Origin and functions of tissue macrophages. *Immunity* **41**, 21–35.
- Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shores, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., et al. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473**, 43–49.
- Feinberg, M.W., Wara, A.K., Cao, Z., Lebedeva, M.A., Rosenbauer, F., Iwasaki, H., Hirai, H., Katz, J.P., Haspel, R.L., Gray, S., et al. (2007). The Kruppel-like factor KLF4 is a critical regulator of monocyte differentiation. *EMBO J.* **26**, 4138–4148.
- Felsenfeld, G., and Groudine, M. (2003). Controlling the double helix. *Nature* **421**, 448–453.
- Garber, M., Yosef, N., Goren, A., Raychowdhury, R., Thielke, A., Guttman, M., Robinson, J., Minie, B., Chevrier, N., Itzhaki, Z., et al. (2012). A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. *Mol. Cell* **47**, 810–822.
- Gautier, E.L., Chow, A., Spanbroek, R., Marcelin, G., Greter, M., Jakubzick, C., Bogunovic, M., Leboeuf, M., van Rooijen, N., Habenicht, A.J., et al. (2012a). Systemic analysis of PPAR γ in mouse macrophage populations reveals marked diversity in expression with critical roles in resolution of inflammation and airway immunity. *J. Immunol.* **189**, 2614–2624.
- Gautier, E.L., Shay, T., Miller, J., Greter, M., Jakubzick, C., Ivanov, S., Helft, J., Chow, A., Elpek, K.G., Gordonov, S., et al.; Immunological Genome Consortium (2012b). Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages. *Nat. Immunol.* **13**, 1118–1128.
- Gautier, E.L., Ivanov, S., Williams, J.W., Huang, S.C., Marcelin, G., Fairfax, K., Wang, P.L., Francis, J.S., Leone, P., Wilson, D.B., et al. (2014). Gata6 regulates aspartoacylase expression in resident peritoneal macrophages and controls their survival. *J. Exp. Med.* **211**, 1525–1531.
- Ghisletti, S., Barozzi, I., Mietton, F., Polletti, S., De Santa, F., Venturini, E., Gregory, L., Lonie, L., Chew, A., Wei, C.-L., et al. (2010). Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. *Immunity* **32**, 317–328.
- Ginhoux, F., and Jung, S. (2014). Monocytes and macrophages: developmental pathways and tissue homeostasis. *Nat. Rev. Immunol.* **14**, 392–404.
- Ginhoux, F., Greter, M., Leboeuf, M., Nandi, S., See, P., Gokhan, S., Mehler, M.F., Conway, S.J., Ng, L.G., Stanley, E.R., et al. (2010). Fate mapping analysis reveals that adult microglia derive from primitive macrophages. *Science* **330**, 841–845.
- Gross, D.S., and Garrard, W.T. (1988). Nuclease hypersensitive sites in chromatin. *Annu. Rev. Biochem.* **57**, 159–197.
- Haldar, M., Kohyama, M., So, A.Y., Kc, W., Wu, X., Briseño, C.G., Satpathy, A.T., Kretzer, N.M., Arase, H., Rajasekaran, N.S., et al. (2014). Heme-mediated SPI-C induction promotes monocyte differentiation into iron-recycling macrophages. *Cell* **156**, 1223–1234.
- Happle, C., Lachmann, N., Skuljec, J., Wetzke, M., Ackermann, M., Brenning, S., Mucci, A., Jirmo, A.C., Groos, S., Mirenska, A., et al. (2014). Pulmonary transplantation of macrophage progenitors as effective and long-lasting therapy for hereditary pulmonary alveolar proteinosis. *Sci. Transl. Med.* **6**, 250ra113.
- Hashimoto, D., Chow, A., Noizat, C., Teo, P., Beasley, M.B., Leboeuf, M., Becker, C.D., See, P., Price, J., Lucas, D., et al. (2013). Tissue-resident macrophages self-maintain locally throughout adult life with minimal contribution from circulating monocytes. *Immunity* **38**, 792–804.
- Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* **39**, 311–318.
- Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W., et al. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112.
- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589.
- Hussell, T., and Bell, T.J. (2014). Alveolar macrophages: plasticity in a tissue-specific context. *Nat. Rev. Immunol.* **14**, 81–93.
- Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Elefant, N., Paul, F., Zaretsky, I., Mildner, A., Cohen, N., Jung, S., Tanay, A., and Amit, I. (2014). Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343**, 776–779.
- Joseph, S.B., Bradley, M.N., Castrillo, A., Bruhn, K.W., Mak, P.A., Pei, L., Hogenesch, J., O'Connell, R.M., Cheng, G., Saez, E., et al. (2004). LXR-dependent gene expression is important for macrophage survival and the innate immune response. *Cell* **119**, 299–309.
- Kohyama, M., Ise, W., Edelson, B.T., Wilker, P.R., Hildner, K., Mejia, C., Frazier, W.A., Murphy, T.L., and Murphy, K.M. (2009). Role for Spi-C in the development of red pulp macrophages and splenic iron homeostasis. *Nature* **457**, 318–321.

- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* *10*, R25.
- Lara-Astiaso, D., Weiner, A., Lorenzo-Vivas, E., Zaretzky, I., Jaitin, D.A., David, E., Keren-Shaul, H., Mildner, A., Winter, D., Jung, S., et al. (2014). Immunogenetics. Chromatin state dynamics during blood formation. *Science* *345*, 943–949.
- Laslo, P., Spooner, C.J., Warmflash, A., Lancki, D.W., Lee, H.J., Sciammas, R., Gantner, B.N., Dinner, A.R., and Singh, H. (2006). Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell* *126*, 755–766.
- Lavin, Y., and Merad, M. (2013). Macrophages: gatekeepers of tissue integrity. *Cancer Immunol. Res.* *1*, 201–209.
- Mortha, A., Chudnovskiy, A., Hashimoto, D., Bogunovic, M., Spencer, S.P., Belkaid, Y., and Merad, M. (2014). Microbiota-dependent crosstalk between macrophages and ILC3 promotes intestinal homeostasis. *Science* *343*, 1249288.
- Murray, P.J., Allen, J.E., Biswas, S.K., Fisher, E.A., Gilroy, D.W., Goerdt, S., Gordon, S., Hamilton, J.A., Ivashkiv, L.B., Lawrence, T., et al. (2014). Macrophage activation and polarization: nomenclature and experimental guidelines. *Immunity* *41*, 14–20.
- Nord, A.S., Blow, M.J., Attanasio, C., Akiyama, J.A., Holt, A., Hosseini, R., Phouanenavong, S., Plajzer-Frick, I., Shoukry, M., Afzal, V., et al. (2013). Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. *Cell* *155*, 1521–1531.
- Okabe, Y., and Medzhitov, R. (2014). Tissue-specific signals control reversible program of localization and functional polarization of macrophages. *Cell* *157*, 832–844.
- Ostuni, R., Piccolo, V., Barozzi, I., Polletti, S., Termanini, A., Bonifacio, S., Curina, A., Prosperini, E., Ghisletti, S., and Natoli, G. (2013). Latent enhancers activated by stimulation in differentiated cells. *Cell* *152*, 157–171.
- Paolicelli, R.C., Bolascho, G., Pagani, F., Maggi, L., Scianni, M., Panzanelli, P., Giustetto, M., Ferreira, T.A., Guiducci, E., Dumas, L., et al. (2011). Synaptic pruning by microglia is necessary for normal brain development. *Science* *333*, 1456–1458.
- Rosas, M., Davies, L.C., Giles, P.J., Liao, C.T., Kharfan, B., Stone, T.C., O'Donnell, V.B., Fraser, D.J., Jones, S.A., and Taylor, P.R. (2014). The transcription factor Gata6 links tissue macrophage phenotype and proliferative renewal. *Science* *344*, 645–648.
- Schafer, D.P., Lehrman, E.K., Kautzman, A.G., Koyama, R., Mardinly, A.R., Yamasaki, R., Ransohoff, R.M., Greenberg, M.E., Barres, B.A., and Stevens, B. (2012). Microglia sculpt postnatal neural circuits in an activity and complement-dependent manner. *Neuron* *74*, 691–705.
- Schneider, C., Nobs, S.P., Kurrer, M., Rehrauer, H., Thiele, C., and Kopf, M. (2014). Induction of the nuclear receptor PPAR- γ by the cytokine GM-CSF is critical for the differentiation of fetal monocytes into alveolar macrophages. *Nat. Immunol.* *15*, 1026–1037.
- Schulz, C., Gomez Perdiguero, E., Chorro, L., Szabo-Rogers, H., Cagnard, N., Kierdorf, K., Prinz, M., Wu, B., Jacobsen, S.E.W., Pollard, J.W., et al. (2012). A lineage of myeloid cells independent of Myb and hematopoietic stem cells. *Science* *336*, 86–90.
- Speliotes, E.K., Kowall, N.W., Shanti, B.F., Kosofsky, B., Finklestein, S.P., and Leifer, D. (1996). Myocyte-specific enhancer binding factor 2C expression in gerbil brain following global cerebral ischemia. *Neuroscience* *70*, 67–77.
- Stergachis, A.B., Neph, S., Reynolds, A., Humbert, R., Miller, B., Paige, S.L., Vernot, B., Cheng, J.B., Thurman, R.E., Sandstrom, R., et al. (2013). Developmental fate and cellular maturity encoded in human regulatory DNA landscapes. *Cell* *154*, 888–903.
- Suzuki, T., Arumugam, P., Sakagami, T., Lachmann, N., Chalk, C., Sallese, A., Abe, S., Trapnell, C., Carey, B., Moritz, T., et al. (2014). Pulmonary macrophage transplantation therapy. *Nature* *514*, 450–454.
- Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* *25*, 1105–1111.
- van Furth, R., Cohn, Z.A., Hirsch, J.G., Humphrey, J.H., Spector, W.G., and Langevoort, H.L. (1972). The mononuclear phagocyte system: a new classification of macrophages, monocytes, and their precursor cells. *Bull. World Health Organ.* *46*, 845–852.
- Varol, C., Vallon-Eberhard, A., Elinav, E., Aychek, T., Shapira, Y., Luche, H., Fehling, H.J., Hardt, W.-D., Shakhar, G., and Jung, S. (2009). Intestinal lamina propria dendritic cell subsets have different origin and functions. *Immunity* *31*, 502–512.
- Virolainen, M. (1968). Hematopoietic origin of macrophages as studied by chromosome markers in mice. *J. Exp. Med.* *127*, 943–952.
- Winter, D.R., and Amit, I. (2014). The role of chromatin dynamics in immune cell development. *Immunol. Rev.* *261*, 9–22.
- Yamanaka, R., Barlow, C., Lekstrom-Himes, J., Castilla, L.H., Liu, P.P., Eckhaus, M., Decker, T., Wynshaw-Boris, A., and Xanthopoulos, K.G. (1997). Impaired granulopoiesis, myelodysplasia, and early lethality in CCAAT/enhancer binding protein epsilon-deficient mice. *Proc. Natl. Acad. Sci. USA* *94*, 13187–13192.
- Yang, C.Y., Chen, J.B., Tsai, T.F., Tsai, Y.C., Tsai, C.Y., Liang, P.H., Hsu, T.L., Wu, C.Y., Netea, M.G., Wong, C.H., and Hsieh, S.L. (2013). CLEC4F is an inducible C-type lectin in F4/80-positive cells and is involved in alpha-galactosylceramide presentation in liver. *PLoS ONE* *8*, e65070.
- Yona, S., Kim, K.-W., Wolf, Y., Mildner, A., Varol, D., Breker, M., Strauss-Ayali, D., Viukov, S., Guillems, M., Misharin, A., et al. (2013). Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis. *Immunity* *38*, 79–91.
- Zigmond, E., Varol, C., Farache, J., Elmali, E., Satpathy, A.T., Friedlander, G., Mack, M., Shpigel, N., Boneca, I.G., Murphy, K.M., et al. (2012). Ly6C hi monocytes in the inflamed colon give rise to proinflammatory effector cells and migratory antigen-presenting cells. *Immunity* *37*, 1076–1090.
- Zigmond, E., Bernshtein, B., Friedlander, G., Walker, C.R., Yona, S., Kim, K.W., Brenner, O., Krauthgamer, R., Varol, C., Müller, W., and Jung, S. (2014). Macrophage-restricted interleukin-10 receptor deficiency, but not IL-10 deficiency, causes severe spontaneous colitis. *Immunity* *40*, 720–733.

Environment Drives Selection and Function of Enhancers Controlling Tissue-Specific Macrophage Identities

David Gosselin,^{1,6} Verena M. Link,^{1,2,6} Casey E. Romanoski,^{1,6} Gregory J. Fonseca,¹ Dawn Z. Eichenfield,¹ Nathanael J. Spann,¹ Joshua D. Stender,¹ Hyun B. Chun,¹ Hannah Garner,^{3,4} Frederic Geissmann,^{3,4} and Christopher K. Glass^{1,5,*}

¹Department of Cellular and Molecular Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0651, USA

²Faculty of Biology, Department II, Ludwig-Maximilians Universität München, Planegg-Martinsried 82152, Germany

³Centre for Molecular and Cellular Biology of Inflammation, King's College London, London SE1 1UL, UK

⁴Peter Gorer Department of Immunobiology, King's College London, London SE1 1UL, UK

⁵Department of Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0651, USA

⁶Co-first author

*Correspondence: ckg@ucsd.edu

<http://dx.doi.org/10.1016/j.cell.2014.11.023>

SUMMARY

Macrophages reside in essentially all tissues of the body and play key roles in innate and adaptive immune responses. Distinct populations of tissue macrophages also acquire context-specific functions that are important for normal tissue homeostasis. To investigate mechanisms responsible for tissue-specific functions, we analyzed the transcriptomes and enhancer landscapes of brain microglia and resident macrophages of the peritoneal cavity. In addition, we exploited natural genetic variation as a genome-wide “mutagenesis” strategy to identify DNA recognition motifs for transcription factors that promote common or subset-specific binding of the macrophage lineage-determining factor PU.1. We find that distinct tissue environments drive divergent programs of gene expression by differentially activating a common enhancer repertoire and by inducing the expression of divergent secondary transcription factors that collaborate with PU.1 to establish tissue-specific enhancers. These findings provide insights into molecular mechanisms by which tissue environment influences macrophage phenotypes that are likely to be broadly applicable to other cell types.

INTRODUCTION

Macrophages are phagocytic cells of the innate immune system that populate every organ, making key contributions to their development, functions, and protection against infections and injuries (Geissmann et al., 2010; Gordon et al., 2014; Wynn et al., 2013). Accordingly, each population of tissue macrophages must adapt to its surrounding environment and engage in tissue-specific functions to be effective auxiliary cells. In sup-

port of this, recent mRNA profiling studies revealed significant differences between distinct populations of resident tissue macrophages (Gautier et al., 2012; Okabe and Medzhitov, 2014). Thus, in spite of common elements shared across all subtypes of tissue macrophages, including dependency on the transcription factor PU.1 and signaling downstream of the CSF1 receptor for ontology and survival (Schulz et al., 2012; Wynn et al., 2013), each subset of tissue macrophage possesses its own unique gene expression profile that presumably allows it to function in synergy with the tissue in which it resides.

Accumulating evidence suggests that signaling factors derived from tissue environments play key roles in promoting the ontology and phenotype of the residing macrophage populations. For example, absence of TGF- β 1 signaling in the mouse brain impairs the development of the microglia population (Butovsky et al., 2014; Makwana et al., 2007). In the peritoneum, omentum-derived retinoic acid (RA) promotes expression of Gata6 in a subpopulation of local macrophages (Okabe and Medzhitov, 2014). Interestingly, Gata6 expression is exclusive to this particular tissue macrophage population, and decreasing or eliminating its expression interferes with their functions and survival (Gautier et al., 2012, 2014; Okabe and Medzhitov, 2014; Rosas et al., 2014).

Precisely how these and other signals act on macrophages at the genomic level to promote specialized phenotypes and unique transcriptional signatures remains unknown. However, strong evidence suggests that enhancers, which are fundamental determinants of gene expression, may play a key role in this context (Andersson et al., 2014; Levine, 2010; Shlyueva et al., 2014). Enhancers, in comparison to promoters, exhibit significant enrichment for combinations of DNA recognition motifs that correspond to binding sites for lineage-determining transcription factors (LDTFs), which are required for the development of distinct cell types. Different patterns of LDTF expression drive the selection of cell-specific repertoires of enhancers that are considered to be central to the establishment of cell identity and regulatory potential.

Studies of primary macrophages and B cells indicated that PU.1 acts as an essential LDTF that contributes to the selection

of a large fraction of the cell-specific enhancer-like elements in each of these cell types (Barozzi et al., 2014; Ghisletti et al., 2010; Heinz et al., 2010). Macrophage-specific enhancer selection by PU.1 required collaborative interactions with additional macrophage-restricted transcription factors (TFs), including C/EBP and AP-1 factors (Heinz et al., 2013). In contrast, B-cell-specific enhancer selection by PU.1 required collaborative interactions with B-cell-restricted factors, including EBF and E2A (Heinz et al., 2010).

Pre-existing enhancer landscapes occupied by PU.1 and/or C/EBP factors were shown to be the major sites that bound signal-dependent transcription factors (SDTFs), such as NF κ B, nuclear receptors, and STAT proteins (Ostuni et al., 2013; Heinz et al., 2010). A similar hierarchical relationship for LDTFs and SDTFs was found in regulatory T cells, embryonic stem cells, and dendritic cells (Mullen et al., 2011; Samstein et al., 2012; Garber et al., 2012). The collaborative and hierarchical relationship of LDTFs and SDTFs at pre-existing enhancers was validated at the level of the DNA template by studies of effects of natural genetic variation on enhancer selection and function (Heinz et al., 2013). Mutations in PU.1 motifs causing loss of PU.1 binding resulted in loss of the collaborative binding of C/EBP α . Conversely, mutations in C/EBP motifs causing loss of C/EBP α binding resulted in a loss of collaborative binding of PU.1. Either type of mutation abolished signal-dependent binding of NF κ B, whereas mutations in NF κ B motifs that abolish NF κ B binding rarely affected the binding of PU.1 or C/EBP α . However, in contrast to the picture at pre-existing enhancers, NF κ B was also shown to be capable of selecting “latent” or “de novo” enhancers by collaborating with PU.1 to bind to genomic locations lacking prior features associated with active enhancers (Kaikkonen et al., 2013; Ostuni et al., 2013). These observations provide an example of an environmentally driven modification of the enhancer repertoire by a broadly expressed SDTF that is nonetheless cell type specific due to the obligatory participation of PU.1.

Given that each tissue environment is distinguished by a unique combination of signaling factors, it is likely that gene expression in each corresponding macrophage population is under the control of distinct combinations of SDTFs that can modulate the activity of a pre-existing enhancer repertoire to achieve context-dependent gene expression. In addition, it is also possible that environmental signals control the expression and activities of TFs that result in selection of tissue-specific enhancers, analogous to the establishment of “latent” or “de novo” enhancers. Here, we sought to determine the extent to which environment shapes distinct macrophage enhancer repertoires and the underlying mechanisms.

RESULTS

Environment-Specific Gene Expression

To investigate mechanisms responsible for tissue-specific macrophage phenotypes, we isolated microglia (MG; brain macrophages) and two distinct populations of resident peritoneal macrophages (RPMs) that are discriminated by cell-surface expression levels of MHCII—large peritoneal macrophages (LPMs, low MHCII) and small peritoneal macrophages (SPMs,

high MHCII)—by flow cytometry (Figures S1A and S1B available online) (Ghosn et al., 2010; Okabe and Medzhitov, 2014) (Figure 1A). These three populations of macrophages allow comparisons of gene expression and epigenetic landscapes in distinct macrophage populations residing in the same environment (i.e., LPMs versus SPMs), as well as different environments (i.e., LPM versus MG). In addition, we included thioglycollate-elicited peritoneal macrophages (TGEMs) and bone-marrow-derived macrophages (BMDMs) for comparison, as these macrophages, although maintained in culture conditions, are widely used models of macrophage biology that are derived from different sources (Figure 1A).

Gene expression profiles determined by RNA sequencing (RNA-seq) from independent biological replicates revealed substantial differences in the patterns of gene expression across the different macrophage populations examined (Figures 1B, 1C, and S1C and Table S1), in agreement with previous studies (Gautier et al., 2012; Okabe and Medzhitov, 2014). In particular, ~7,000 genes are differently expressed in MG compared to LPMs (p value < 0.01), with >500 genes being >16-fold more highly expressed in MG and >600 genes being >16-fold more highly expressed in LPMs. On the other hand, LPMs and SPMs share strong similarities (Figure 1C), with SPMs expressing only 108 genes > 16-fold higher than LPMs, and LPMs expressing only 5 genes > 16-fold higher than SPMs. These results corroborate many previous findings, including the highest level of expression of *Cx3cr1* in MG and the selective expression of *Gata6* in RPMs (Figure 1D) (Cardona et al., 2006; Gautier et al., 2012; Jung et al., 2000; Okabe and Medzhitov, 2014). Interestingly, *Ciita*, a transcription factor that regulates MHCII expression (Steimle et al., 1993), is preferably expressed in the SPM population (Figure 1D). Finally, gene clustering analyses confirmed that, whereas LPMs and SPMs show highly similar gene expression, MG differ substantially from the other macrophage subsets (Figure 1E). TGEMs and BMDMs are also more similar to one another than either one is to any of the three in vivo subsets, potentially reflecting the similarity of the cell culture environment. Overall, these findings suggest a strong role of environment in determining macrophage gene expression.

Common and Distinct Macrophage Enhancer Repertoires

The dissimilarities in gene expression between different macrophage subsets revealed by RNA-seq analysis imply important differences in how these cells organize and/or use their enhancer repertoires. To examine this, we analyzed dimethylation status of lysine 4 of histone 3 (H3K4me2) and acetylation status of lysine 27 of histone H3 (H3K27ac) by chromatin immunoprecipitation sequencing (ChIP-seq) in these cells (Figure S2 and Tables S2, S3, and S4). H3K4me2 marks promoters and enhancers (He et al., 2010; Kaikkonen et al., 2013), whereas H3K27ac correlates positively with transcriptional activity at these elements (Creyghton et al., 2010). Deposition of H3K4me2 results from the binding of LDTFs and other TFs but is not necessarily associated with enhancer activity. We therefore use a heuristic of defining H3K4me2-positive/H3K27ac-negative regions as “primed” and regions positive for both marks as “active.” Genomic annotation enabled segregation of these regions into

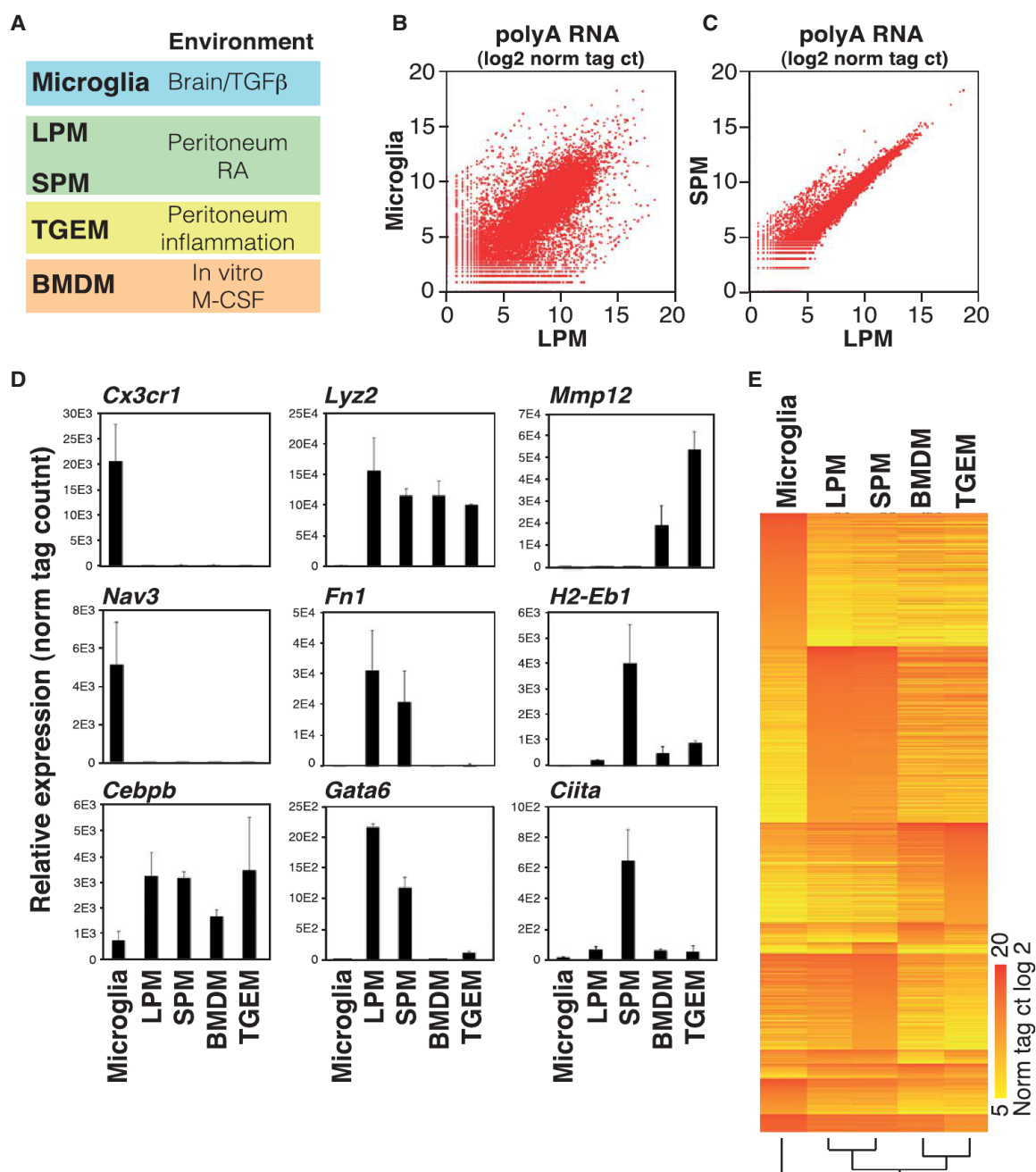


Figure 1. Variation in Gene Expression in Different Macrophage Subsets

(A) Macrophage subsets used for analysis and corresponding environmental factors (see Figures S1A and S1B for sorting protocols). (B and C) Scatterplots illustrating relative gene expression of polyA-selected RNA transcripts in MG compared to LPMs (B) and SPMs compared to LPMs (C). Values are log2 of tag counts normalized to 10^7 uniquely mapped tags. See Figure S1C for a representative replicate. (D) Relative gene expression means for the indicated genes are shown from replicate RNA-seq experiments (error bars represent SD). (E) Heat map of transcripts exhibiting an expression value of at least 64 normalized tags in at least one subset and differing in expression by at least 16-fold in at least one of the indicated subsets. See also Table S1.

promoters or enhancers by proximity to gene transcriptional start sites (TSS). Notably, the pattern of H3K4me2 deposition in MG substantially differs from that of LPMs (Figure 2A), indicating selection of distinct regulatory landscapes. Of 7,937

promoters marked by H3K4me2 in one or both subsets, 275 exhibit >4-fold differences (3%), far fewer than the ~1,700 mRNAs exhibiting >16-fold differences in expression. Of 36,607 regions > 500 bp from TSS marked by H3K4me2 in one

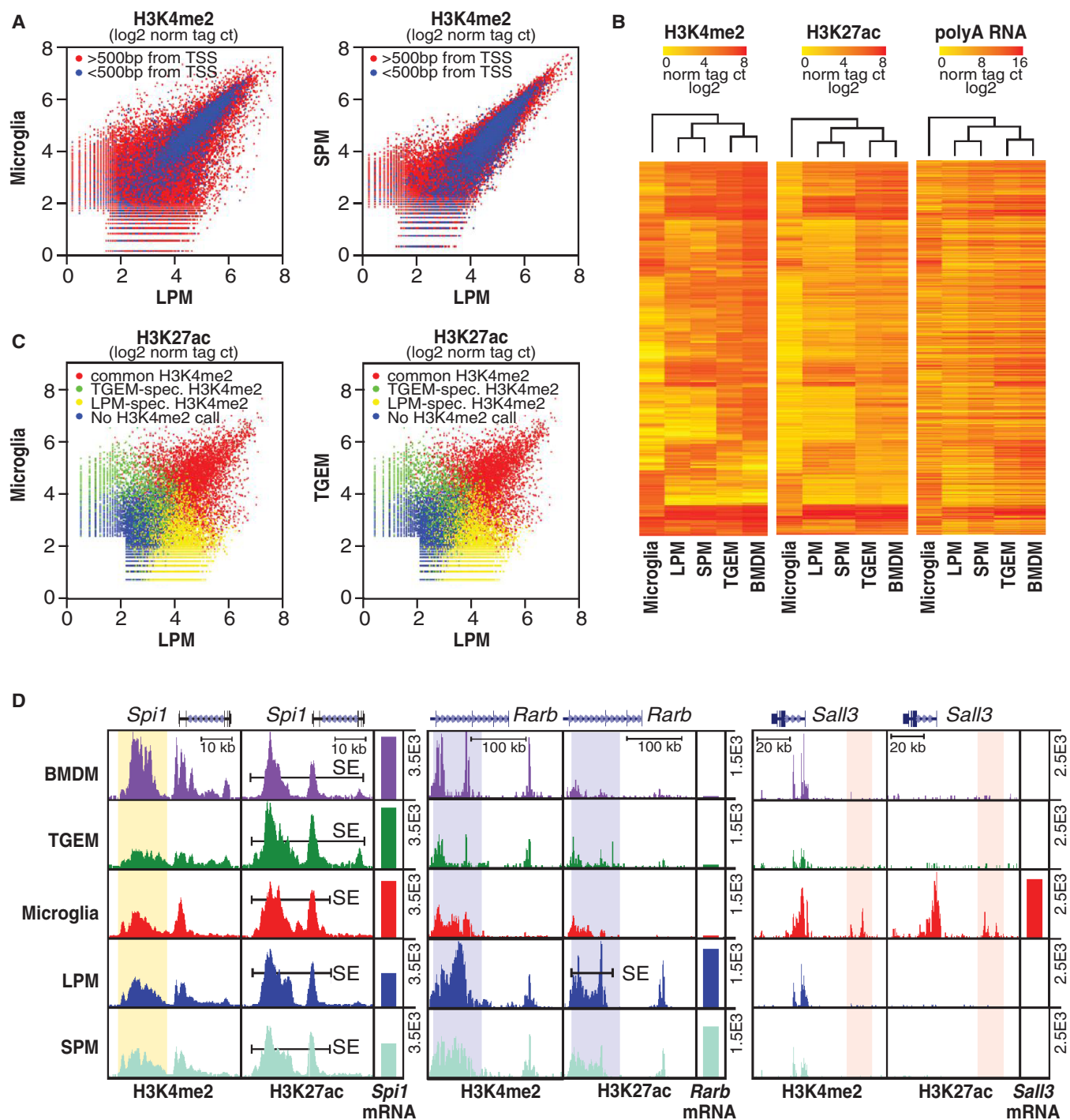


Figure 2. Variation in Enhancer Landscapes in Different Macrophage Subsets

(A) Scatterplots of normalized H3K4me2 tag counts at genomic regions marked by significant H3K4me2 tags in LPMs and/or MG (left) or LPMs and/or SPMs (right). Points colored in blue are within 500 bp of a TSS. See Figure S2 for representative replicates.

(B) Heatmaps of normalized H3K4me2, H3K27ac, and nearest expressed gene RNA-seq tag counts at genomic locations showing >4-fold pairwise differences in H3K4me2 tag counts between at least two of the five macrophage subtypes. Row order is the same for all three data types.

(C) Scatterplots of normalized H3K27ac tag counts at genomic regions marked by significant H3K27ac tags in LPMs and/or MG (left) or LPMs and/or SPMs (right). Points are colored red if genomic locations are also marked by H3K4me2 (>16 tags) in both subsets, green if marked by H3K4me2 selectively in MG (left) or SPMs (right), yellow if marked by H3K4me2 selectively in LPMs, or blue if not associated with H3K4me2 in either subset.

(legend continued on next page)

or both subsets, 9,083 exhibit >4-fold differences (24%). The vast majority of differential H3K4me2-marked regions are thus distant from promoters and correspond to potential enhancers. In contrast to the comparison of LPMs and MG, both the enhancer and promoter repertoires of the two subsets of RPMs share a much higher degree of similarity (Figure 2A). Furthermore, clustering analyses of the H3K4me2 deposition pattern revealed that MG were more divergent from the other subsets than any two other macrophage subsets are from one another, which is consistent with gene expression data (Figure 2B).

H3K27ac was present at a large fraction of H3K4me2-marked regions and generally but imperfectly correlated with nearest gene expression (Figure 2B). Overlap of the H3K27ac data with H3K4me2-defined enhancers allowed the identification of common but quantitatively differently activated enhancers, as well as activation of enhancers unique to one subset. Figure 2C illustrates such comparisons for LPMs versus MG and LPMs versus TGEMs. Genomic regions marked by H3K4me2 in both subsets are color coded in red and represent activation of an enhancer landscape that is primed in both subsets. In contrast, regions exclusively marked by H3K4me2 in LPMs, shown in yellow, represent LPM-specific enhancers. Conversely, regions exclusively marked by H3K4me2 in MG or TGEMs, indicated in green, represent MG or TGEM-specific enhancers, respectively. Comparing LPMs versus MG, 60% of the active enhancers resided at common regions of H3K4me2, 30% at LPM-specific regions, and 10% at MG-specific regions. Specific examples are indicated in Figure 2D. As expected, the *Spi1* enhancer, controlling expression of PU.1, is marked by H3K4me2 and H3K27ac in all macrophage populations. Interestingly, the RA-inducible *Rarb* gene is also marked by H3K4me2 in all macrophage populations, but high H3K27ac is only observed in LPMs and SPMs, suggesting a role of local RA in enhancer activation. Finally, the *Sal13* gene, which is exclusively and highly expressed in MG, is near a genomic region that is exclusively marked by H3K4me2 and H3K27ac in MG. In sum, these analyses provide strong evidence that both differential activation of a common enhancer landscape and the selection of subset-type-specific enhancers contribute to the specific transcriptional signature of each subset of macrophages.

Tissue-Specific Super-Enhancers Emerge from Common Enhancer Landscapes

Genome-wide analysis of features of active enhancers, including the presence of Mediator and deposition of H3K27ac, indicates marked variation in their local distribution patterns. In all cell types evaluated thus far, ~400–800 regions, representing a small fraction of the genome, exhibit a disproportionately high density of active regulative marks and transcription factor binding (Hnisz et al., 2013; Lovén et al., 2013; Whyte et al., 2013). These regions, recently termed super-enhancers (SEs), are selected in a cell-specific manner and frequently occur near or encompass

genes that play essential roles in defining the identity and function of the corresponding cell type (Hnisz et al., 2013). Although LDTFs are enriched in and likely determine cell-specific SE selection, evidence also suggests that the extracellular environment can influence formation of SEs in endothelial cells (Brown et al., 2014). To investigate this relationship in tissue macrophages, we defined SEs in each macrophage subset based on H3K27ac ChIP-seq. In agreement with previous studies, we observed common and subset-specific SEs, with ~600 to 750 SEs being identified among the five cell types examined. Clustering of these SEs results in the same relationships between subsets as observed using RNA-seq, H3K4me2, or H3K27ac data (Figure 3A). This analysis also revealed a high concordance between the distribution of SEs genome wide and the expression level of the nearest genes (Figure 3A). This strong relationship is further illustrated for SEs and nearest gene expression in MG and LPMs, in which the correlation coefficient was 0.62 (Figure 3B), much higher than that observed for the individual enhancer elements not associated with SE regions in these subsets. This may be due to a more accurate assignment of SEs to their target genes than conventional enhancers.

Approximately 40% to 50% of the SEs in a particular macrophage subset are unique to that subset, illustrated by the Venn diagram of LPM, MG, and TGEM in Figure 3C. In concert with previous findings (Hnisz et al., 2013; Whyte et al., 2013), common SEs are associated with numerous genes important to macrophage ontology and functions, including *Spi1*, *Cebpa*, members of the Irf family, *Csf1r*, *Fcgr2b*, *Ctsb*, etc. (Figure 3D). This pattern is exemplified by the region upstream of *Spi1*, which is scored as a SE in all five subsets (Figure 2D). In contrast, many SEs are macrophage subset specific and reside near or surround genes that are highly differentially expressed (Figure 3E). Although some SEs exhibit highly specific H3K4me2 and H3K27ac markings, such as the LPM-specific SE upstream of *Gata6* (Figure 3E), the majority of SEs are located at regions that are marked by H3K4me2 in multiple macrophage subsets but only attain SE status in one or a few subsets. For example, LPM-specific SEs reside in the vicinity of *Rarb* (Figure 2D) and *Alox15* (Figure 3E) genes, which are selectively expressed in LPMs but that also exhibit H3K4me2 in other macrophage subsets. Similar relationships are observed for the MG-specific SEs surrounding *Gpr56* and *Cx3cr1* and the TGEM-specific SEs surrounding *Fabp5* and *Gpnmb* (Figure 3E). These findings suggest that environmental signals play roles in the transition of collections of primed enhancers to genomic regions exhibiting features of SEs.

PU.1 Colocalizes with Distinct TF Motifs at Subset-Specific Enhancers

The observation that PU.1 localization to macrophage- or B-cell-specific enhancers is dependent on collaborative interactions with alternate LDTFs (Heinz et al., 2010) led us to consider the possibility that an assessment of PU.1 binding in different macrophage subsets might yield insights into the TFs that

(D) UCSC browser images of selected genomic regions with corresponding RNA-seq data plotted as bar graphs. Bars labeled SE indicate super-enhancers, and vertical highlights designate regions of interest for subset-common (*Spi1*) or subset-specific (*Rarb* and *Sal13*) loci. All data are normalized to input and library dimension.

See also Tables S2, S3, and S4.

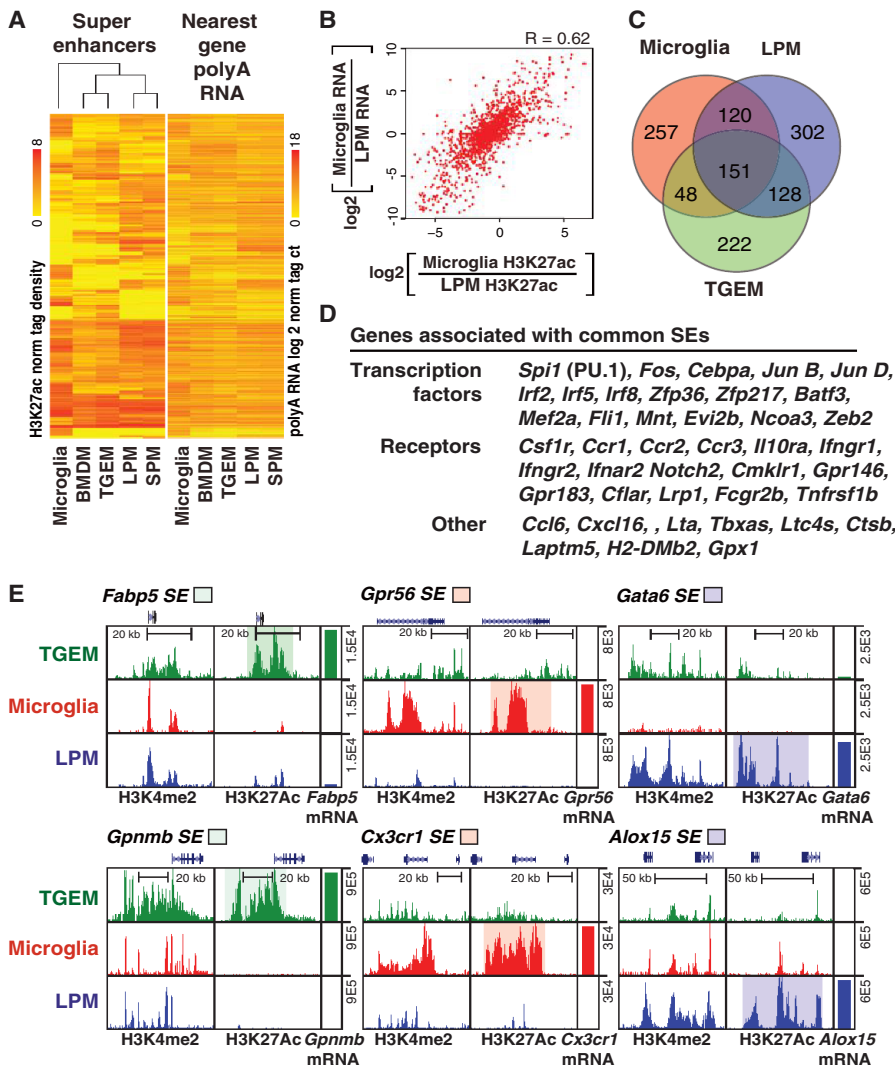


Figure 3. Variation in Super-Enhancer Landscapes in Different Macrophage Subsets

(A) Heatmaps of H3K27ac tag densities at super-enhancers and RNA-seq tag densities at nearest genes. Rows are ordered the same for both plots. (B) Scatterplot of the relationship between ratio of MG to LPMs H3K27ac tag density at super-enhancers (x axis) and the ratio of nearest gene expression (y axis). (C) Venn diagram indicating overlap and specificity of super-enhancers in MG, LPMs, and TGMs. (D) Examples of genes associated with common super-enhancers. (E) UCSC genome browser images of selected subset-specific super-enhancers and associated genes with subset-specific regions of interest highlighted.

IRF, KLF, and GATA transcription factor family members (Figure 4C). Conversely, MG-specific PU.1-binding sequences were coenriched for a PU.1-IRF composite sequence and motifs corresponding to CTCFL, HIC2, MEF2, and SMAD TFs (Figure 4D). In addition, by using alternative subset-specific PU.1-binding sites as background, motifs recognized by retinoic acid receptors (e.g., NR2F2) were identified to be coenriched with PU.1-binding sites in LPMs (Figure 4C).

Previous studies indicated that motifs for collaborative binding partners of PU.1 typically reside within ~100 bp of the PU.1 motif itself (Barozzi et al., 2014; Heinz et al., 2010). We therefore analyzed the genomic distance distribution of enriched motifs (from Figures 4C and 4D) within a 400 bp window relative to the

drive the selection of subset-specific enhancers. We therefore extended existing genome-wide binding profiles for PU.1 to include MG, LPMs, and SPMs. These studies indicated that PU.1 bound to both common and subset-specific genomic locations, exemplified for LPMs and MG in Figure 4A (all comparisons in Tables S2 and S3). The great majority of subset-specific binding sites were observed at distal regions (>500 bp from an mRNA TSS, Figure 4A), which is consistent with the patterns of H3K4me2 (Figure 2A). Examples of LPM-specific and MG-specific binding sites for PU.1 in enhancer-like regions vicinal to *Msr1* (expressed exclusively in LPMs) and *Nav2* (expressed exclusively in MG) genes are illustrated in Figure 4B.

De novo motif enrichment analysis of 200 bp sequences encompassing PU.1 peaks identified the identical PU.1 recognition motif in both LPMs and MG as the most enriched sequence. However, completely different motifs were coenriched within the two subsets (Figures 4C and 4D). Using GC content-matched genomic sequence as background, enriched sequences specific to LPMs corresponded to motifs known to bind C/EBP, AP-1,

bound PU.1 motif of LPM- and MG-specific PU.1 peak sets (Figure 4E). This analysis indicated that C/EBP, AP1, and GATA motifs frequently occurred near PU.1-bound motifs in LPMs, but not in MG, indicating that genomic loci containing PU.1 and closely spaced C/EBP, AP-1, or GATA motifs were more likely to become LPM-specific enhancers. The GATA motif was selectively enriched in LPMs relative to MG, suggesting a fundamental difference for the LPM resident population compared to elicited macrophages (Figure 4E). In contrast, the SMAD motif showed MG specificity (Figure 4E), which is consistent with TGF β signaling in the brain. These findings provide evidence that selection of subset-specific enhancers is in part driven by collaborative interactions between PU.1 and alternative sets of TFs in each subset.

Use of Natural Genetic Variation to Validate and Discover Collaborative TFs

Although motif enrichment suggests the identities of TFs that contribute to the function of subset-specific enhancers, this

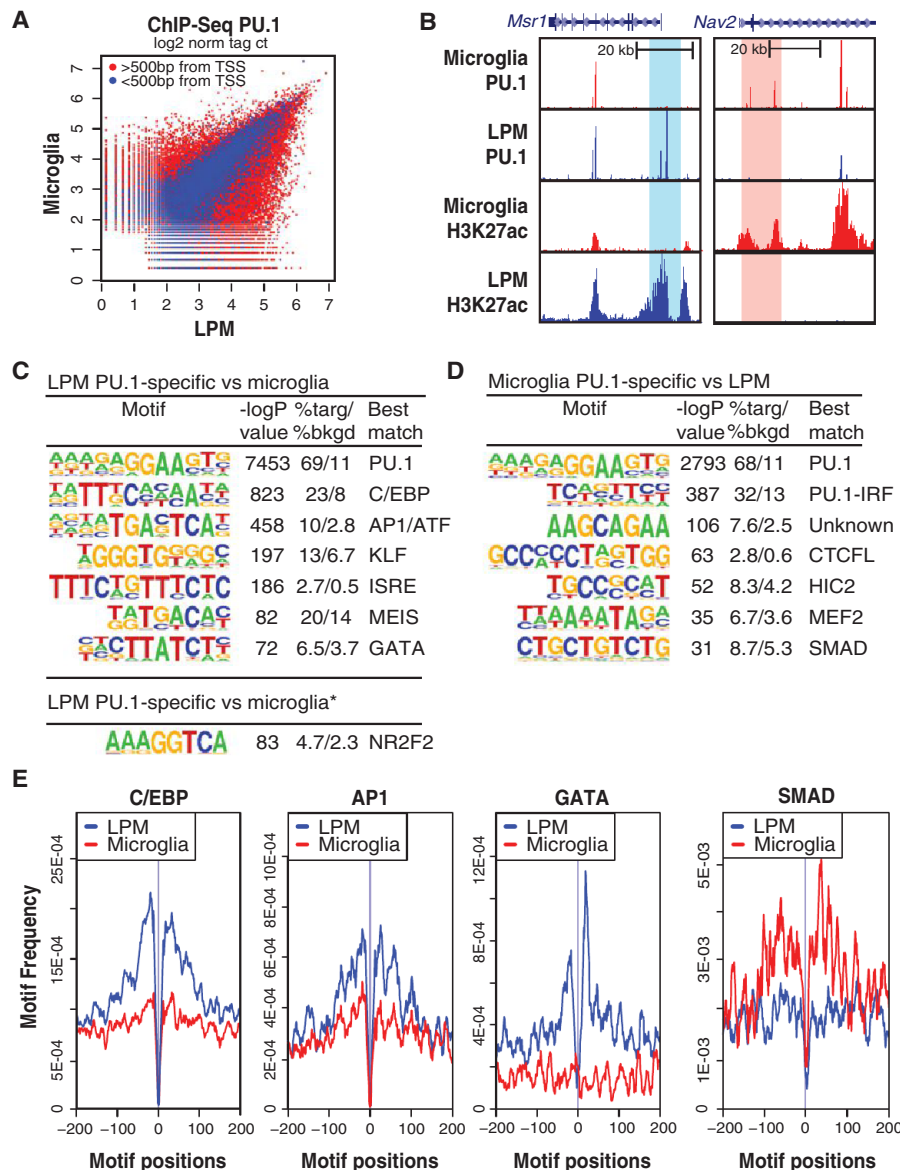


Figure 4. PU.1 Binds to Subset-Specific Enhancers

(A) Scatterplot of normalized tag counts for PU.1 peaks in MG versus LPMs. Points colored blue are within 500 bp of the TSS.

(B) UCSC genome browser images of PU.1 binding in the vicinity of the *Msr1* and *Nav2* genes in MG and LPMs cells and association with H3K27ac highlighting specific regions.

(C) Motifs enriched in the vicinity of PU.1-binding sites that are specific for LPMs versus MG using a random GC-corrected genomic background (top) or a background corresponding to MG-specific PU.1 peaks (bottom).

(D) Motifs enriched in the vicinity of PU.1-binding sites that are specific for MG using a random GC-corrected genomic background.

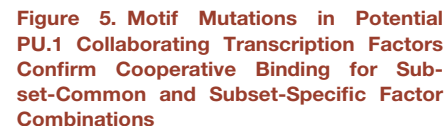
(E) Distribution plots of motif frequencies (y axis) for the indicated motifs within 400bp centered on the PU.1 motif at genomic loci bound specifically by PU.1 in LPMs (blue) or MG (red).

Compared to C57BL/6J (C57) mice, NOD mice have about 5 million SNPs and indels, whereas SPRET mice have about 40 million (Keane et al., 2011). This variation is associated with corresponding levels of strain-specific binding of PU.1, illustrated for LPMs derived from C57 and SPRET mice (Figure 5A). Similar observations are made with respect to MG (Table S6). Approximately 8-fold fewer strain-specific PU.1-binding sites were identified in LPMs and MG derived from NOD mice compared to C57, which is consistent with the lower number of variants between these two strains. Strain-specific binding of PU.1 was associated with corresponding strain-specific H3K4me2 and H3K27ac marks (Figure 5B), suggesting that many strain-specific PU.1-binding sites localize to functional enhancers.

approach does not establish whether or not they are required for collaborative binding. Loss-of-function strategies are challenging for this purpose because many of the identified motifs are recognized by multiple members of corresponding TF families. An alternative means to test for collaborative binding is to mutate motifs recognized by the TF family of interest and determine whether this results in loss of binding of a nearby factor. We considered the possibility that this could be accomplished for informative motifs on a genome-wide scale by leveraging the vast degree of natural genetic variation provided by inbred laboratory and wild strains of mice.

To explore the potential of this approach to validate and discover TFs required for collaborative binding and function of PU.1, we determined the genome-wide patterns of PU.1, H3K4me2, and H3K27ac in LPMs and MG isolated from NOD/ShiLtJ (NOD) and SPRET/EiJ (SPRET) mice (Table S5).

To search for motifs mediating DNA binding by collaborative TFs, we analyzed strain-specific binding of PU.1 that was not associated with mutations in PU.1 recognition motifs. This was accomplished by scanning a 200 bp window surrounding PU.1-binding sites lacking PU.1 motif mutations for the presence of the DNA recognition motifs of the 100 most highly expressed TFs in LPMs and MG in C57 or the alternate (NOD or SPRET) genomic sequence. Mutated loci were then queried for a corresponding decrease in PU.1 binding relative to the unmutated strain. The significant result for ISRE motif mutations affecting PU.1 binding in LPMs is exemplified in Figure 5C. The ISRE was found to be mutated in the vicinity of PU.1-binding sites 93 times in LPMs isolated from C57 mice (indicated by red hash lines in Figure 5C) and 106 times in LPMs isolated from SPRET mice (indicated by blue hash lines in Figure 5C). PU.1 binding strength is rank ordered from most C57 specific at left



(E) Heatmap showing p values resulting from analysis described in (C) and (D) for motif mutations best matching transcription factors indicated on x axis. Each motif was tested for affecting PU.1 binding between C57 and NOD and between C57 and SPRET both in MG and LPMs (y axis).

To investigate the importance of tissue environment in maintenance of specific macrophage phenotypes, we placed LPMs and MG into culture under the influence of IL-34 or M-CSF for 7 days. Whereas M-CSF is important to peritoneal macrophages (Witmer-Pack et al., 1993), IL-34 is critical for proper MG

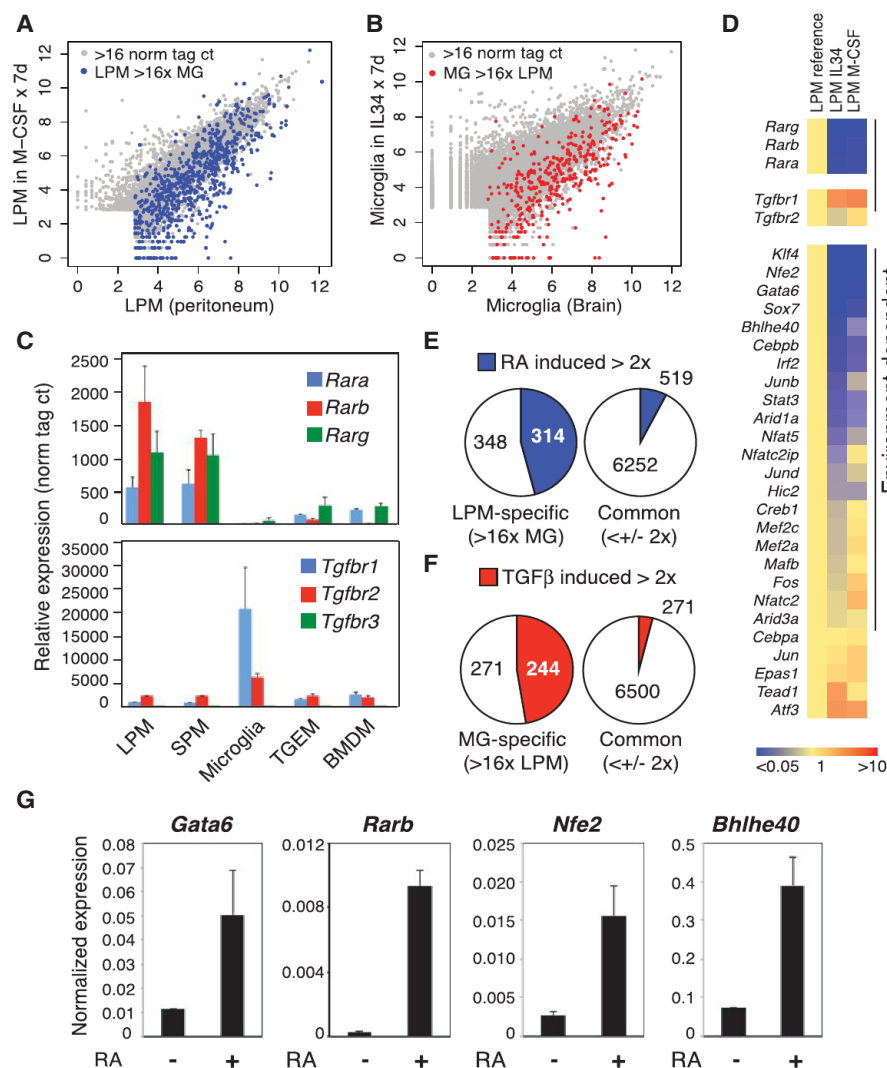


Figure 6. Environmental Influence on Gene Expression in LPMs and Microglia

(A and B) Scatterplots illustrating relative gene expression of RNA transcripts in freshly isolated LPMs compared to LPMs maintained in culture for 7 days (A) and freshly isolated MG compared to MG in culture for 7 days (B). Genes specific to LPMs are colored blue in (A) and specific to MG are red in (B).

(C) Normalized gene expression values for members of the RAR and TGFβ receptor family members.

(D) Heatmap showing the fold-change of RNAs for the indicated transcription factors upon removal from the peritoneal cavity and culture with IL-34 or M-CSF.

(E and F) Effects of chronic stimulation with RA in M-CSF and/or IL34 (E) on LPM-specific or common mRNAs or TGFβ in M-CSF or IL34 (F) on MG-specific or common mRNAs.

(G) qPCR validation of maintained expression by RA of key transcription factors in cultured LPMs (error bars indicate SD). See also Figure S4.

in each macrophage subset. The mRNAs encoding all three RA receptors (*Rara*, *Rarb*, and *Rarg*) are highly and selectively expressed in LPMs and SPMs, whereas mRNAs encoding the TGFβ receptors *Tgfb1* and *Tgfb2* are preferentially expressed in MG (Figure 6C). Interestingly, expression of all three retinoic acid receptors is markedly reduced when LPMs are placed into culture in the presence of M-CSF or IL-34, whereas the expression of *Tgfb1* is markedly increased under these conditions (Figure 6D). Thus, environment

ontology and/or survival in vivo (Greter et al., 2012; Wang et al., 2012). This environmental transition resulted in vast changes in gene expression (Table S7). Comparison of the gene expression program of LPMs freshly purified from the peritoneal cavity with LPMs maintained in M-CSF for 7 days is illustrated in Figure 6A. Data points colored in blue represent genes that are expressed more than 16-fold higher in LPMs than MG, indicating that the LPM-specific program of gene expression is preferentially lost in culture. Comparison of the gene expression program of MG freshly isolated from the brain or maintained in culture in the presence of IL-34 for 7 days is illustrated in Figure 6B. Data points colored in red represent genes that are expressed more than 16-fold higher in MG than LPMs, indicating that the MG-specific program of gene expression is preferentially lost in culture. In both LPMs and MG, many genes exhibiting low levels of expression in vivo are markedly upregulated in culture.

In view of recent findings indicating important roles of TGFβ signaling in MG and RA signaling in peritoneal macrophages, we examined the expression of the main receptors for these factors

controls the expression of genes responsible for responses to environment-specific signals.

To investigate the extent to which RA and TGFβ influence subset-specific patterns of gene expression, we treated LPMs with RA or TGFβ for 7 days and performed RNA-seq analysis. RA treatment induced expression of nearly half of the LPM-specific genes by more than 2-fold, while inducing about 8% of genes expressed at similar levels in LPMs and MG (Figure 6E). Conversely, nearly 50% of the genes induced more than 2-fold by TGFβ in LPMs in culture are preferentially expressed by MG in vivo, whereas only 4% of the genes expressed at similar levels in LPMs and MG were induced by TGFβ in LPMs (Figure 6F). Thus, RA and TGFβ disproportionately regulate genes that specify LPM and MG-specific phenotypes, respectively.

We next evaluated the expression of TFs that recognize motifs identified as putative binding sites for collaborative partners of PU.1 in LPMs through analysis of strain-specific PU.1 binding. Remarkably, expression of the majority of TFs best matched to motifs identified by strains analysis was environment dependent (Figure 6D). A similar pattern was observed when considering all

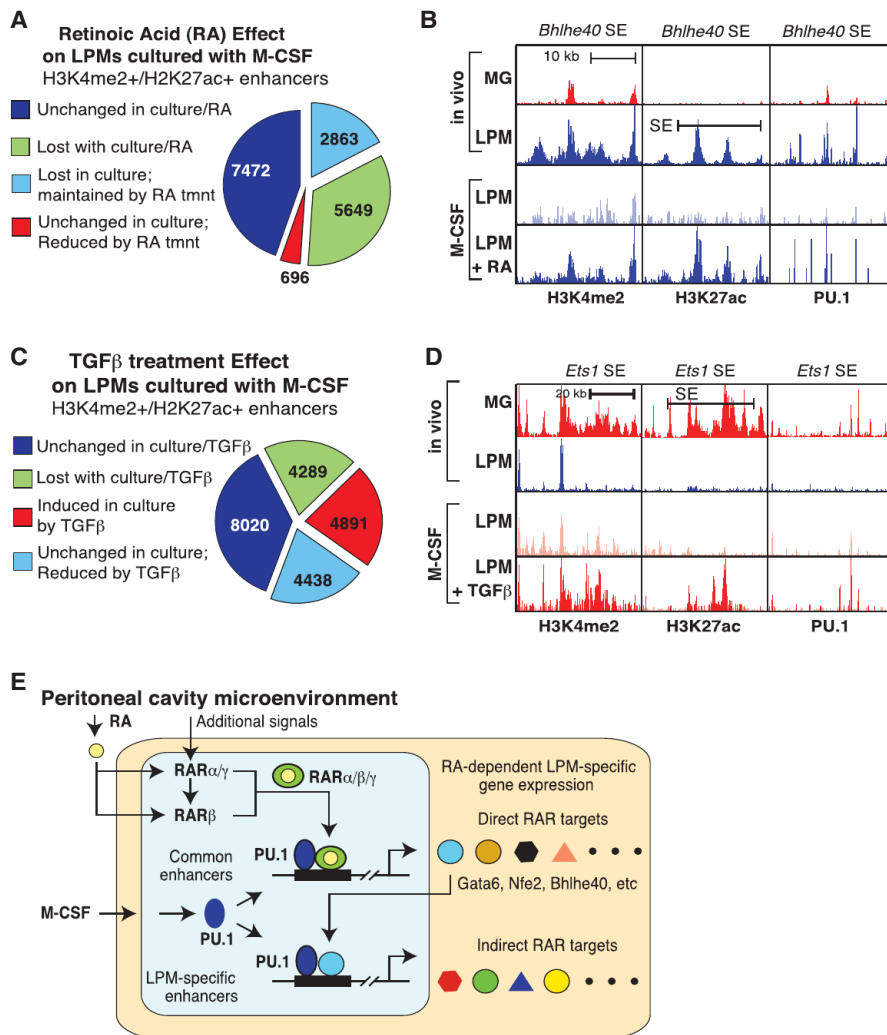


Figure 7. Environmental Influence on Enhancer Landscapes in LPMs and Microglia

(A) Effects of culture environment and RA chronic stimulation on the enhancer landscape of LPMs. (B) UCSC browser images displaying effects of culture environment and RA chronic stimulation on H3K4me2, H3K27ac, and PU.1 binding at the *Bhlhe40* locus in LPMs. (C) Effects of culture environment and chronic stimulation with TGF β on the enhancer landscape of LPMs. (D) UCSC browser images displaying effects of culture environment and chronic stimulation with TGF- β 1 on H3K4me2, H3K27ac, and PU.1 binding at the *Ets1* locus in LPMs. (E) Hierarchical model for mechanisms by which the peritoneal environment induces the enhancer landscape and gene expression signature of LPMs. See Discussion for details. See also Table S7.

the enhancer-like regions (Figure 7A). One-third of these lost enhancer elements were maintained by RA treatment (Figure 7A). Of the 302 LPM-specific SEs identified in Figure 3, 223 (74%) no longer met SE criteria, indicating a disproportionate sensitivity to loss of environmental signals. This pattern is exemplified by the SE associated with *Bhlhe40*, which, in addition to substantial reduction in the histone signature of enhancers, also exhibits reduced PU.1 binding (Figure 7B). Notably, H3K4me2, H3K27ac, and PU.1 binding are largely maintained by RA treatment.

Maintenance of LPMs in M-CSF plus TGF β resulted in marked changes in the

LPM enhancer landscape in comparison to culture in M-CSF alone, which is consistent with the preferential effects of TGF β on a MG-specific program of gene expression (Figure 7C). Treatment with TGF β increased the enhancer signature by more than 2-fold at ~25% of pre-existing enhancers. Conversely, TGF β reduced enhancer signatures at ~25% enhancer-like elements that were stable upon transfer to culture in M-CSF. Induced enhancers are exemplified by a genomic region in the vicinity of the *Ets1* gene (Figure 7D) that is preferentially expressed in MG and is highly induced by TGF β . Culture of LPMs in M-CSF results in appearance of PU.1 binding and H3K4me2 modification, with TGF β treatment leading to substantial increases in H3K27ac and a marked increase in gene expression.

DISCUSSION

Mechanisms Underlying Tissue-Specific Enhancer Selection and Activation

The present studies provide evidence for a hierarchical model in which the distinct environments of the brain and peritoneal cavity differentially activate a common set of primed enhancers and

members of each TF family capable of recognizing these motifs (Figure S4A). RNA-seq analysis further suggested that several of these factors were inducible by RA. This response was confirmed under M-CSF treatment conditions for *Gata6* and *Rarb*, consistent with previous studies (Okabe and Medzhitov, 2014), as well as for *Bhlhe40* and *Nfe2* (Figures 6G and S4B). However, expression of *Rara*, *Rarg*, and most of the other factors illustrated in Figure 6C was not RA inducible. Thus, the environment modulates in LPMs the expression of collaborative and SDTFs through both RA-dependent and RA-independent mechanisms.

Hierarchical Effects of Environment on Macrophage Enhancer Landscapes

To gain insights into mechanisms underlying effects of environment on macrophage gene expression, we performed ChIP-seq analysis for H3K4me2, H3K27ac, and PU.1 in LPMs maintained in M-CSF and the presence or absence of RA or TGF β for 7 days. Transition of LPMs from the peritoneal cavity to a tissue culture environment containing M-CSF led to a >2-fold reduction in H3K4me2 and/or H3K27ac at approximately half of

their target genes that, in turn, promote the selection and activation of subset-specific enhancer repertoires. The combinatorial activation of both common and subset-specific enhancers enables context-dependent regulation of genes required for specialized functions of MG and RPMs. Aspects of this model as they pertain to the RA-dependent program of gene expression specific to peritoneal macrophages are illustrated in Figure 7E. Common to all macrophage subsets, stimulation of signaling pathways downstream of the M-CSF receptor by M-CSF and/or IL-34, which are present in the environment in a largely tissue-non-specific manner, ensures survival and promotes PU.1 expression (Sarrazin et al., 2009). PU.1 is a critical LDTFs required for all macrophage subsets that functions to select common and cell-specific enhancers through collaborative interactions with other TFs. These regions of PU.1 binding in turn serve as subset-specific sites of action for various types of SDTFs.

Within the peritoneal cavity, environment-specific signals control the expression and activities of TFs that act upon primed enhancers that are common to multiple macrophage subsets (Figure 7E). A particularly important signal is omentum-derived RA, which has been shown to be essential for development and function of LPMs through its activation of RAR β and induction of *Gata6* (Okabe and Medzhitov, 2014). We find that all three high-affinity retinoic acid receptor genes (*Rara*, *Rarb*, and *Rarg*) are preferentially expressed in the peritoneal cavity and that this expression requires continual maintenance by the peritoneal cavity environment. However, only *Rarb* expression is preserved by RA treatment in culture, indicating that expression of *Rara* and *Rarg* is under the control of as-yet-unidentified factors. We speculate that the expression of RAR α and RAR γ is necessary for full induction of RAR β expression in response to environmental RA and that this positive feedback loop is important for amplification of the RA signal and activation of direct RA target genes. These findings imply that at least two environmental signals are required for initiating the RA-dependent peritoneal macrophage phenotype, one being RA itself and the second being a signal or signals required for RAR α and RAR γ expression.

Activated retinoic acid receptors primarily function as SDTFs that act at a common set of primed enhancers established by PU.1 and other LDTFs that are expressed across macrophage subsets. Importantly, direct RA target genes include *Gata6*, *Bhlhe40*, and *Nfe2*, which were identified as putative interacting partners of PU.1 through analysis of effects of natural genetic variation. We propose that RAR-dependent induction of these factors results in collaborative interactions with PU.1 that drive environment-specific selection of LPM-specific enhancers (Figure 7E).

Of note, *Gata6*, *Bhlhe40*, and *Nfe2*, as well as all three retinoic acid receptors, reside in or near peritoneal macrophage-specific SEs that are lost when LPMs are removed from the peritoneal cavity. Our findings suggest that an analogous hierarchy operates in MG, driven in part by TGF β signaling and SMAD TFs. Although the present studies have focused on PU.1, we expect that additional macrophage LDTFs function in an analogous manner to set up macrophage-specific, PU.1-independent enhancers.

Use of Natural Genetic Variation to Validate and Discover Collaborative TFs

Here, we demonstrate the use of the natural genetic variation provided by inbred strains of mice as a powerful means to validate and discover collaborative TFs. By measuring strain-specific binding of PU.1 in macrophages derived from genetically diverse strains of mice, we identified motifs for several different classes of TFs in which strain-specific mutations were highly correlated with the loss of binding of PU.1 to nonmutated PU.1 recognition motifs. Interestingly, the expression of a significant fraction of the TFs recognizing these motifs is dependent on environment.

Many of the motifs identified by analysis of strain-specific binding of PU.1 are recognized by TFs that have well-established roles in macrophage biology. Some, such as C/EBP α and C/EBP β , are documented to function as factors that enable collaborative binding of PU.1 in macrophages (Heinz et al., 2010), supporting the validity of the approach. Although the biological role of *Gata6* in the development and function of LPMs is established (Okabe and Medzhitov, 2014; Rosas et al., 2014), the present studies suggest that a key molecular function of *Gata6* is to collaborate with PU.1, and likely other macrophage LDTFs, to drive the selection of LPM-specific enhancers. *Bhlhe40* and *Nfe2* represent examples of putative collaborative partners of PU.1 that have not as yet been linked to macrophage-specific functions. *Bhlhe40*, also known as *Dec1*, *Stra13*, and *Sharp2*, has previously been shown to be inducible by RA and to act as both as a repressor and activator (Boudjelal et al., 1997; Ivanova et al., 2004), raising the possibility that it could contribute to selection of LPM-specific enhancers, as well as suppress genes that become active when LPMs are removed from the peritoneal cavity. *Nfe2* is a bZip transcription factor that is broadly expressed in the hematopoietic system and has been established to play important roles in erythropoiesis and megakaryocyte development (Andrews, 1998). The present findings provide a rationale for further investigation of roles of *Bhlhe40*, *Nfe2*, and other TFs identified as putative collaborative binding partners of PU.1.

The use of natural genetic variation as a strategy for identification of TFs required for enhancer selection can in principle be applied to any cell type in which ChIP-seq can be performed for an index LDTF. In addition, although not a focus of the present studies, the variation in enhancer selection and activity observed in macrophages derived from different inbred strains of mice was associated with strain-specific differences in LPM and MG gene expression. Such changes in gene expression are presumably linked to both molecular phenotypes such as eQTLs and to the marked phenotypic differences exhibited by these mice that are influenced by tissue resident macrophage populations, such as relative susceptibility or resistance to metabolic, cardiovascular, infectious, and neurodegenerative diseases (Civelek and Lusis, 2014; Threadgill and Churchill, 2012). The principle of collaborative binding, which serves as the basis for the motif discovery method described here, is directly applicable to investigating mechanisms by which noncoding variants may exert phenotypic effects in a cell-type-specific and/or context-dependent manner. In concert, these approaches enable insights into gene-by-environment interactions and the genetic architecture of molecular and complex disease traits.

Tuning Enhancer Landscapes and Gene Expression to Context-Specific Functions

The present studies reveal that each macrophage subset uniquely possesses a distinct set of active enhancers, including subset-specific SEs, which are associated with strong preferential expression of nearby genes. In LPMs, for example, which populate a very potent immunogenic environment, *Gbp2b* and *Alox15* are associated with SE activity, and we note that the protein products of these genes are critical regulators of immunity, in particular inflammation and tolerance (Pilla et al., 2014; Uderhardt et al., 2012; Yamamoto et al., 2012). In contrast to LPMs, MG reside in the immune-privileged environment of the brain. As with LPMs, however, our observations suggest that MG adopt a unique phenotype that is again strongly contributed by distinct enhancers and SEs to accomplish tissue-specific functions required for brain homeostasis. For example, SEs in MG include genomic loci associated with the *Cx3cr1* and *Gpr56* genes, among others. Interestingly, both genes are highly relevant to brain functions, regulating synaptic pruning and efficient cortical patterning during brain development (Paolicelli et al., 2011; Piao et al., 2004). Together, our studies reveal an intricate relationship between the organization of the genome of tissue macrophage and their surrounding environment.

Divergent Macrophage Gene Expression in a Common Environment

Distinct macrophage populations can coexist in a similar environment, as illustrated by the copresence of LPMs and SPMs in the peritoneum. Although these cells are highly concordant with respect to gene expression and organization of their enhancer landscapes, consistent with exposure to common tissue-derived signals, strong points of divergence can nonetheless discriminate the two. These observations raise the possibility that differences in origin and ontology play important roles in determining these later-stage differences (Perdiguero et al., 2014; Schulz et al., 2012). Thus, the impact of developmental history on the regulation of enhancer repertoires and gene expression of different tissue macrophages remains a fundamental open question to be addressed in future studies.

EXPERIMENTAL PROCEDURES

Mice

Seven-week-old C57BL/6J, NOD/ShiLtJ, and SPRET/EiJ male mice were purchased from Jackson Labs and used at 8 to 9 weeks of age. All animal procedures were in accordance with University of California, San Diego research guidelines for the care and use of laboratory animals.

Microglia Isolation

Mice were anaesthetized with CO₂ and quickly perfused intracardially with ice-cold DPBS. Whole brains were removed and gently mechanically homogenized on ice. Cells were fractionated by Percoll gradient centrifugation, and microglia-enriched fractions were further purified by cell sorting according to the scheme described in Figure S1A and Extended Experimental Procedures.

Peritoneal Macrophage Isolation

Following euthanization, peritoneal cells were collected by lavage of the peritoneum with ice-cold staining buffer. LPM and SPM subsets were purified based on relative expression of MHCII and other markers described in Figure S1B and Extended Experimental Procedures.

Thioglycollate-Elicited and Bone-Marrow-Derived Macrophages Cultures

TGEMs were harvested by peritoneal lavage with 20 ml ice-cold PBS 4 days after peritoneal injection of 3 ml Thioglycollate broth. Both TGEMs and BMDMs were cultured as described in Heinz et al. (2010). See also Extended Experimental Procedures.

ChIP-Seq

Macrophages were fixed at room temperature with 1% paraformaldehyde/PBS containing 1 mM sodium butyrate for 10 min and quenched with glycine. 2.0×10^5 to 1.0×10^6 cells were used for ChIP, and samples were processed as previously described (Heinz et al., 2010), with minor modifications noted in the Extended Experimental Procedures. Sequencing libraries were prepared as previously described (Heinz et al., 2010).

RNA Isolation

For RNA-seq, TRIzol (Life Technologies) isolated RNA was either PolyA-selected (MicroPoly(A) Purist kit, Ambion) or subjected to RiboZero rRNA removal (Epicenter).

Quantitative PCR, RNA-Seq Library Preparation, and Sequencing

Libraries for RNA sequencing were generated as previously described (Heinz et al., 2013). See Extended Experimental Procedures for details and qRT-PCR primer sequences.

Data Analysis

Fastq files from sequencing experiments were mapped to individual genomes for the mouse strain of origin using default parameters for STAR (Dobin et al., 2013) (RNA-seq) and Bowtie2 (Langmead and Salzberg, 2012) (ChIP-seq). NOD/ShiLtJ and SPRET/EiJ custom genomes were generated from invariant positions of the mm10 sequence with alleles replaced by those reported in VCF files from the Mouse Genomes Project (Keane et al. 2011). Mapped data were analyzed with HOMER (Heinz et al., 2010), custom R, and Perl scripts.

ACCESSION NUMBERS

Raw and processed data are provided in the Gene Expression Omnibus (GEO) under accession number GSE62826.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, four figures, and seven tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2014.11.023>.

AUTHOR CONTRIBUTIONS

D.G. and C.K.G. designed the study. D.G., G.J.F., D.Z.E., N.J.S., J.D.S., H.B.C., and H.G. performed experiments. D.G., V.L., C.E.R., F.G., and C.K.G. analyzed and interpreted the data. D.G., V.L., C.E.R., and C.K.G. wrote the manuscript.

ACKNOWLEDGMENTS

The authors would like to thank Lynn Bautista, Andrea Crotti, and Renee F. Glass for valuable assistance with preparation of the manuscript and Sven Heinz for assistance with the ChIP-seq library preparation protocol. The authors would also like to thank the UCSD Human Embryonic Stem Cell Core Facility for assistance with cell sorting. These studies were primarily supported by NIH grants DK091183, CA17390, and DK063491 and the San Diego Center for Systems Biology (GM085764). D.G. was supported by a Canadian Institutes of Health Research Fellowship. C.E.R. was supported by the American Heart Association (12POST11760017) and the NIH Pathway to Independence Award (1K99HL12348). D.Z.E. was supported by an American Heart Association predoctoral fellowship (12PRE11610007). F.G. was supported by a Wellcome Trust Senior Investigator award (WT101853MA).

Received: November 2, 2014
 Revised: November 16, 2014
 Accepted: November 17, 2014
 Published: December 4, 2014

REFERENCES

- Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., et al.; FANTOM Consortium (2014). An atlas of active enhancers across human cell types and tissues. *Nature* 507, 455–461.
- Andrews, N.C. (1998). The NF-E2 transcription factor. *Int. J. Biochem. Cell Biol.* 30, 429–432.
- Barozzi, I., Simonatto, M., Bonifacio, S., Yang, L., Rohs, R., Ghisletti, S., and Natoli, G. (2014). Coregulation of transcription factor binding and nucleosome occupancy through DNA features of mammalian enhancers. *Mol. Cell* 54, 844–857.
- Boudjelal, M., Taneja, R., Matsubara, S., Bouillet, P., Dolle, P., and Chambon, P. (1997). Overexpression of Stra13, a novel retinoic acid-inducible gene of the basic helix-loop-helix family, inhibits mesodermal and promotes neuronal differentiation of P19 cells. *Genes Dev.* 11, 2052–2065.
- Brown, J.D., Lin, C.Y., Duan, Q., Griffin, G., Federation, A.J., Paranal, R.M., Bair, S., Newton, G., Lichtman, A.H., Kung, A.L., et al. (2014). NF- κ B Directs Dynamic Super Enhancer Formation in Inflammation and Atherogenesis. *Mol. Cell* 56, 219–231.
- Butovsky, O., Jedrychowski, M.P., Moore, C.S., Cialic, R., Lanser, A.J., Gabriely, G., Koeglsparger, T., Dake, B., Wu, P.M., Doykan, C.E., et al. (2014). Identification of a unique TGF- β -dependent molecular and functional signature in microglia. *Nat. Neurosci.* 17, 131–143.
- Cardona, A.E., Pioro, E.P., Sasse, M.E., Kostenko, V., Cardona, S.M., Dijkstra, I.M., Huang, D., Kidd, G., Dombrowski, S., Dutta, R., et al. (2006). Control of microglial neurotoxicity by the fractalkine receptor. *Nat. Neurosci.* 9, 917–924.
- Civelek, M., and Lusis, A.J. (2014). Systems genetics approaches to understand complex traits. *Nat. Rev. Genet.* 15, 34–48.
- Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. USA* 107, 21931–21936.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.
- Garber, M., Yosef, N., Goren, A., Raychowdhury, R., Thielke, A., Guttman, M., Robinson, J., Minie, B., Chevrier, N., Itzhaki, Z., et al. (2012). A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. *Mol. Cell* 47, 810–822.
- Gautier, E.L., Shay, T., Miller, J., Greter, M., Jakubzick, C., Ivanov, S., Helft, J., Chow, A., Elpek, K.G., Gordonov, S., et al.; Immunological Genome Consortium (2012). Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages. *Nat. Immunol.* 13, 1118–1128.
- Gautier, E.L., Ivanov, S., Williams, J.W., Huang, S.C., Marcelin, G., Fairfax, K., Wang, P.L., Francis, J.S., Leone, P., Wilson, D.B., et al. (2014). Gata6 regulates aspartoacylase expression in resident peritoneal macrophages and controls their survival. *J. Exp. Med.* 211, 1525–1531.
- Geissmann, F., Manz, M.G., Jung, S., Sieweke, M.H., Merad, M., and Ley, K. (2010). Development of monocytes, macrophages, and dendritic cells. *Science* 327, 656–661.
- Ghisletti, S., Barozzi, I., Mietton, F., Polletti, S., De Santa, F., Venturini, E., Gregory, L., Lonie, L., Chew, A., Wei, C.L., et al. (2010). Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. *Immunity* 32, 317–328.
- Ghosn, E.E., Cassado, A.A., Govoni, G.R., Fukuhara, T., Yang, Y., Monack, D.M., Bortoluci, K.R., Almeida, S.R., Herzenberg, L.A., and Herzenberg, L.A. (2010). Two physically, functionally, and developmentally distinct peritoneal macrophage subsets. *Proc. Natl. Acad. Sci. USA* 107, 2568–2573.
- Gordon, S., Pluddemann, A., and Martinez Estrada, F. (2014). Macrophage heterogeneity in tissues: phenotypic diversity and functions. *Immunol. Rev.* 262, 36–55.
- Greter, M., Lelios, I., Pelczar, P., Hoeffel, G., Price, J., Leboeuf, M., Kündig, T.M., Frei, K., Ginhoux, F., Merad, M., and Becher, B. (2012). Stroma-derived interleukin-34 controls the development and maintenance of langerhans cells and the maintenance of microglia. *Immunity* 37, 1050–1060.
- He, H.H., Meyer, C.A., Shin, H., Bailey, S.T., Wei, G., Wang, Q., Zhang, Y., Xu, K., Ni, M., Lupien, M., et al. (2010). Nucleosome dynamics define transcriptional enhancers. *Nat. Genet.* 42, 343–347.
- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* 38, 576–589.
- Heinz, S., Romanoski, C.E., Benner, C., Allison, K.A., Kaikkonen, M.U., Orzoco, L.D., and Glass, C.K. (2013). Effect of natural genetic variation on enhancer selection and function. *Nature* 503, 487–492.
- Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. *Cell* 155, 934–947.
- Ivanova, A.V., Ivanov, S.V., Zhang, X., Ivanov, V.N., Timofeeva, O.A., and Lerman, M.I. (2004). STRA13 interacts with STAT3 and modulates transcription of STAT3-dependent targets. *J. Mol. Biol.* 340, 641–653.
- Jung, S., Aliberti, J., Graemmel, P., Sunshine, M.J., Kreutzberg, G.W., Sher, A., and Littman, D.R. (2000). Analysis of fractalkine receptor CX3CR1 function by targeted deletion and green fluorescent protein reporter gene insertion. *Mol. Cell Biol.* 20, 4106–4114.
- Kaikkonen, M.U., Spann, N.J., Heinz, S., Romanoski, C.E., Allison, K.A., Stender, J.D., Chun, H.B., Tough, D.F., Prinjha, R.K., Benner, C., and Glass, C.K. (2013). Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Mol. Cell* 51, 310–325.
- Keane, T.M., Goodstadt, L., Danecek, P., White, M.A., Wong, K., Yalcin, B., Heger, A., Agam, A., Slater, G., Goodson, M., et al. (2011). Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* 477, 289–294.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.
- Levine, M. (2010). Transcriptional enhancers in animal development and evolution. *Curr. Bio.* 20, R754–R763.
- Lovén, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* 153, 320–334.
- Makwana, M., Jones, L.L., Cuthill, D., Heuer, H., Bohatschek, M., Hristova, M., Friedrichsen, S., Ormsby, I., Bueringer, D., Koppius, A., et al. (2007). Endogenous transforming growth factor beta 1 suppresses inflammation and promotes survival in adult CNS. *J. Neurosci.* 27, 11201–11213.
- Mullen, A.C., Orlando, D.A., Newman, J.J., Lovén, J., Kumar, R.M., Bilodeau, S., Reddy, J., Guenther, M.G., DeKoter, R.P., and Young, R.A. (2011). Master transcription factors determine cell-type-specific responses to TGF- β signaling. *Cell* 147, 565–576.
- Okabe, Y., and Medzhitov, R. (2014). Tissue-specific signals control reversible program of localization and functional polarization of macrophages. *Cell* 157, 832–844.
- Ostuni, R., Piccolo, V., Barozzi, I., Polletti, S., Termanini, A., Bonifacio, S., Curina, A., Prosperini, E., Ghisletti, S., and Natoli, G. (2013). Latent enhancers activated by stimulation in differentiated cells. *Cell* 152, 157–171.
- Paolicelli, R.C., Bolasco, G., Pagani, F., Maggi, L., Scianni, M., Panzanelli, P., Giustetto, M., Ferreira, T.A., Guiducci, E., Dumas, L., et al. (2011). Synaptic pruning by microglia is necessary for normal brain development. *Science* 333, 1456–1458.

- Perdiguerro, E.G., Klapproth, K., Schulz, C., Busch, K., Azzoni, E., Crozet, L., Garner, H., Trouillet, C., de Bruijn, M.F., Geissmann, F., et al. (2014). Tissue-resident macrophages originate from yolk sac-derived erythro-myeloid progenitors. *Nature*. Published online December 3, 2014. <http://dx.doi.org/10.1038/nature13989>.
- Piao, X., Hill, R.S., Bodell, A., Chang, B.S., Basel-Vanagaite, L., Straussberg, R., Dobyns, W.B., Qasrawi, B., Winter, R.M., Innes, A.M., et al. (2004). G protein-coupled receptor-dependent development of human frontal cortex. *Science* 303, 2033–2036.
- Pilla, D.M., Hagar, J.A., Haldar, A.K., Mason, A.K., Degrandi, D., Pfeffer, K., Ernst, R.K., Yamamoto, M., Miao, E.A., and Coers, J. (2014). Guanylate binding proteins promote caspase-11-dependent pyroptosis in response to cytoplasmic LPS. *Proc. Natl. Acad. Sci. USA* 111, 6046–6051.
- Rosas, M., Davies, L.C., Giles, P.J., Liao, C.T., Kharfan, B., Stone, T.C., O'Donnell, V.B., Fraser, D.J., Jones, S.A., and Taylor, P.R. (2014). The transcription factor Gata6 links tissue macrophage phenotype and proliferative renewal. *Science* 344, 645–648.
- Samstein, R.M., Arvey, A., Josefowicz, S.Z., Peng, X., Reynolds, A., Sandstrom, R., Neph, S., Sabo, P., Kim, J.M., Liao, W., et al. (2012). Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification. *Cell* 151, 153–166.
- Sarrazin, S., Mossadegh-Keller, N., Fukao, T., Aziz, A., Mourcin, F., Vanhille, L., Kelly Modis, L., Kastner, P., Chan, S., Duprez, E., et al. (2009). MafB restricts M-CSF-dependent myeloid commitment divisions of hematopoietic stem cells. *Cell* 138, 300–313.
- Schulz, C., Gomez Perdiguerro, E., Chorro, L., Szabo-Rogers, H., Cagnard, N., Kierdorf, K., Prinz, M., Wu, B., Jacobsen, S.E., Pollard, J.W., et al. (2012). A lineage of myeloid cells independent of Myb and hematopoietic stem cells. *Science* 336, 86–90.
- Shlyueva, D., Stampfel, G., and Stark, A. (2014). Transcriptional enhancers: from properties to genome-wide predictions. *Nat. Rev. Genet.* 15, 272–286.
- Steimle, V., Otten, L.A., Zufferey, M., and Mach, B. (1993). Complementation cloning of an MHC class II transactivator mutated in hereditary MHC class II deficiency (or bare lymphocyte syndrome). *Cell* 75, 135–146.
- Threadgill, D.W., and Churchill, G.A. (2012). Ten years of the collaborative cross. *G3 (Bethesda)* 2, 153–156.
- Uderhardt, S., Herrmann, M., Oskolkova, O.V., Aschermann, S., Bicker, W., Ipseiz, N., Sarter, K., Frey, B., Rothe, T., Voll, R., et al. (2012). 12/15-lipoxygenase orchestrates the clearance of apoptotic cells and maintains immunologic tolerance. *Immunity* 36, 834–846.
- Wang, Y., Szretter, K.J., Vermi, W., Gilfillan, S., Rossini, C., Cella, M., Barrow, A.D., Diamond, M.S., and Colonna, M. (2012). IL-34 is a tissue-restricted ligand of CSF1R required for the development of Langerhans cells and microglia. *Nat. Immunol.* 13, 753–760.
- Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153, 307–319.
- Witmer-Pack, M.D., Hughes, D.A., Schuler, G., Lawson, L., McWilliam, A., Inaba, K., Steinman, R.M., and Gordon, S. (1993). Identification of macrophages and dendritic cells in the osteopetrotic (op/op) mouse. *J. Cell Sci.* 104, 1021–1029.
- Wynn, T.A., Chawla, A., and Pollard, J.W. (2013). Macrophage biology in development, homeostasis and disease. *Nature* 496, 445–455.
- Yamamoto, M., Okuyama, M., Ma, J.S., Kimura, T., Kamiyama, N., Saiga, H., Ohshima, J., Sasai, M., Kayama, H., Okamoto, T., et al. (2012). A cluster of interferon- γ -inducible p65 GTPases plays a critical role in host defense against *Toxoplasma gondii*. *Immunity* 37, 302–313.

Species-wide Genetic Incompatibility Analysis Identifies Immune Genes as Hot Spots of Deleterious Epistasis

Eunyoung Chae,¹ Kirsten Bomblies,^{1,3} Sang-Tae Kim,^{1,4} Darya Karelina,^{1,2} Maricris Zaidem,¹ Stephan Ossowski,^{1,5} Carmen Martín-Pizarro,¹ Roosa A.E. Laitinen,^{1,6} Beth A. Rowan,¹ Hezi Tenenboim,^{1,6} Sarah Lechner,^{1,7} Monika Demar,¹ Anette Habring-Müller,¹ Christa Lanz,¹ Gunnar Rättsch,^{2,8} and Detlef Weigel^{1,*}

¹Department of Molecular Biology, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany

²Friedrich Miescher Laboratory, Max Planck Society, 72076 Tübingen, Germany

³Present address: Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

⁴Present address: Center for Genome Engineering, Institute for Basic Science, 305-811 Daejeon, South Korea

⁵Present address: Genomic and Epigenomic Variation in Disease Group, Centre for Genomic Regulation and Universitat Pompeu Fabra, 08003 Barcelona, Spain

⁶Present address: Max Planck Institute of Molecular Plant Physiology, 14476 Potsdam-Golm, Germany

⁷Present address: CeGAT GmbH, 72076 Tübingen, Germany

⁸Present address: Computational Biology Center, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

*Correspondence: weigel@weigelworld.org

<http://dx.doi.org/10.1016/j.cell.2014.10.049>

SUMMARY

Intraspecific genetic incompatibilities prevent the assembly of specific alleles into single genotypes and influence genome- and species-wide patterns of sequence variation. A common incompatibility in plants is hybrid necrosis, characterized by autoimmune responses due to epistatic interactions between natural genetic variants. By systematically testing thousands of F₁ hybrids of *Arabidopsis thaliana* strains, we identified a small number of incompatibility hot spots in the genome, often in regions densely populated by nucleotide-binding domain and leucine-rich repeat (NLR) immune receptor genes. In several cases, these immune receptor loci interact with each other, suggestive of conflict within the immune system. A particularly dangerous locus is a highly variable cluster of NLR genes, *DM2*, which causes multiple independent incompatibilities with genes that encode a range of biochemical functions, including NLRs. Our findings suggest that deleterious interactions of immune receptors limit the combinations of favorable disease resistance alleles accessible to plant genomes.

INTRODUCTION

When independently diverging genomes meet in hybrids, the differences that have accumulated over evolutionary time can have detrimental consequences. The ensuing incompatibilities were formally described by Bateson, Dobzhansky, and Muller, who proposed a scenario under which complementary changes occur in two different populations (Bateson, 1909; Dobzhansky, 1937; Muller, 1942); the individual changes are

innocuous in their native genomic contexts, and they reduce viability or fertility only when combined in hybrids (Coyne and Orr, 2004). This type of deleterious, or negative, epistasis has been most prominently studied in interspecific crosses, where the interacting alleles are fixed in the different populations (Maheshwari and Barbash, 2011; Presgraves, 2010; Rieseberg and Blackman, 2010). More recent work has begun to focus on deleterious epistasis within species, where the interacting alleles are polymorphic and segregate in a single intermating population (Corbett-Detig et al., 2013; Hou et al., 2014; Seidel et al., 2008). One can envision a series of evolutionary forces responsible for the emergence of interacting alleles. On the one hand, genetic drift could play a role, with segregating alleles that are neutral or merely mildly deleterious on their own giving rise to synthetic deleterious interactions (Bikard et al., 2009; Phillips and Johnson, 1998). At the other end of the spectrum, adaptation and intragenomic conflicts have been implicated as factors that may contribute to a reduction in hybrid viability or fertility (Crespi and Nosil, 2013; Cutter, 2012; Lachance and True, 2010). Regardless of the primary cause, high levels of sequence divergence at incompatibility loci appear to be positively correlated with deleterious interactions. Ultimately, as lineages diverge and become genetically more differentiated, segregating variants may rise to a high frequency in specific populations and thereby reduce gene flow among them.

Genes of the immune system are particularly polymorphic in many organisms, because of an evolutionary arms race between hosts and pathogens (Quintana-Murci and Clark, 2013; Sackton et al., 2007; Vilches and Parham, 2002). This is also true for plants. Prominent, highly variable components of the plant immune system are the nucleotide-binding domain and leucine-rich repeat (NLR) proteins, with plant genomes often encoding hundreds of NLRs (Cao et al., 2011; Jacob et al., 2013). Plant NLRs typically function as immune receptors that confer disease resistance by monitoring the integrity of important plant proteins or the presence of pathogen effector proteins (Collier and Moffett, 2009).

Apart from having to keep up with the ongoing evolution of individual pathogens, hosts must also accumulate resistance against as many different pathogens as possible. This in turn entails its own dangers, in the form of genetic variants that lead to enhanced immunity but at the same time reduce growth or health due to autoimmunity (Todesco et al., 2010; Trowsdale and Knight, 2013). In plants, especially severe autoimmune phenotypes have been observed in hybrids descended from phenotypically normal parents. This syndrome, hybrid necrosis, is found in intra- and interspecific crosses and spans a range of severity, from cases where only a small subset of F_2 progeny is weakly affected to ones in which all F_1 hybrids die. The lesions and reduced growth of necrotic hybrids are often alleviated at higher temperature, greatly facilitating the genetic analysis of severe cases (Bomblies and Weigel, 2007).

To date, four hybrid necrosis cases due to two-locus epistasis have been studied at the molecular level in tomato, lettuce, and rice. Of the six causal loci that have been positively identified, one encodes an NLR and another a known NLR interactor. In addition, the mapping interval for one of the remaining loci includes several NLRs (Chen et al., 2014; Jeuken et al., 2009; Krüger et al., 2002; Yamamoto et al., 2010). Similarly, the first hybrid necrosis gene positively identified in *Arabidopsis thaliana*, *DANGEROUS MIX1* (*DM1*), encodes an NLR. It interacts with the *DM2* locus, which was mapped to a polymorphic cluster of *RPP1* NLR genes that is probably also responsible for an independent F_2 incompatibility (Alcázar et al., 2009; Bomblies et al., 2007). In natural accessions of *A. thaliana*, the *RPP* subfamily of NLR genes is particularly variable, both in sequence and copy number. The high diversity of *RPP* loci, which recognize different strains of the oomycete pathogen *Hyaloperonospora arabidopsidis* ex *parasitica* (*Hpa*) in an allele-specific manner, points to these genes as actors in an active coevolutionary tug of war between host and pathogen (Allen et al., 2004; Bakker et al., 2006; Holub and Beynon, 1997).

Although there is increasing evidence for natural variation in immunity potentially leading to genetic incompatibilities in plants, species-wide patterns of immune-related deleterious epistasis remain unknown. For example, are specific immune loci especially likely to be involved in deleterious epistasis? Do they interact more often with other immune loci than with nonimmune genes? And is deleterious epistasis correlated with geographic and genetic distance? To systematically investigate which factors in the plant immune system contribute to intraspecific incompatibility, we have examined F_1 progeny from thousands of *A. thaliana* crosses, including all combinations among 80 accessions that represent much of the common genetic diversity in the species (Cao et al., 2011). We have mapped several hybrid necrosis loci to regions of the genome that contain multiple NLR genes in tandem arrays, with different allelic variants at *DM2/RPP1* being responsible for several independent incompatibilities. We also found cases where different alleles at a locus interact with each other, or where independent pairs of alleles at two loci are incompatible with each other. Because many interactions are between components of the plant immune system, we propose that epistatic fitness effects limit the extent to which favorable immune alleles can be combined.

RESULTS

A Systematic Resource for the Discovery of Genetic Incompatibilities

About 2% of crosses between randomly chosen accessions of *A. thaliana* suffer from F_1 hybrid necrosis when grown at 16°C (Bomblies et al., 2007). To determine the incidence of hybrid necrosis and other F_1 weakness syndromes more systematically and to obtain insights into how genetic kinship affects the probability of hybrid necrosis, we turned to 80 accessions that had their genomes sequenced in the first phase of the 1001 Genomes Project (Cao et al., 2011). These 80 accessions represent much of the common diversity across the species' native range, and can thus provide insights into the distribution of hybrid necrosis alleles in both the global and local populations. To facilitate the large number of crosses, male sterile lines were derived by knocking down the floral homeotic gene *AP3* (Wuest et al., 2012). Together with additional crosses that mostly involved accessions carrying known *DM* alleles, we assembled a total of 6,409 crosses. This collection comprised 3,330 unique parental combinations; 3,160 of these made up a complete half-diallel of the 80 accessions (Table S1 available online).

The most common morphological defects seen in F_1 hybrids at 16°C were dwarfism and tissue necrosis, which fell into five classes of increasing severity, including two new extreme classes (Figure 1; Figure S1). In the previously described cases, morphological defects largely disappeared at 23°C (Bomblies et al., 2007). Class 4 phenotypes were only suppressed at 28°C, whereas class 5 individuals died as seedlings at all temperatures tested. Most necrotic hybrids had only mild defects (103 cases in class 1), 29 cases were intermediate (classes 2–4), and 10 cases were not genetically tractable due to lethality (class 5). Our threshold for identifying necrosis was quite stringent, and there might well be many more weak cases.

Mapping and Identification of Incompatibility Loci

From the 142 F_1 hybrid necrosis cases described here and the 27 identified previously (Bomblies et al., 2007), we chose 7 for further genetic analysis. In addition to obvious phenotypes, we prioritized cases where at least one of the causal alleles was likely to be present in multiple genetic backgrounds, as judged by one parent producing similar F_1 phenotypes with several other parents. Thus, the selected cases are likely to represent 31 of the 48 intermediate, genetically tractable hybrid necrosis examples in our collection.

The fraction of affected individuals in F_2 populations indicated that hybrid phenotypes in five cases were due to pairwise interactions between genetically separable loci (Table S2). Segregation ratios in *Ey1.5-2* × *ICE228* and for the *Bla-1* × *Hh-0* lesioning trait were consistent with effects arising from heterozygous disadvantage at single regions of the genome (Table S2). We mapped causal loci mostly using genotyping by sequencing (GBS) of individual F_2 plants (Elshire et al., 2011; Poland et al., 2012) and quantitative trait locus (QTL) methods (Figures 2A–2H; Table S3). For leaf twisting in *Bla-1* × *Hh-0*, we used whole-genome sequencing of pooled DNA from F_1 -like individuals segregating in selfed backcrossed hybrid (BC) 5 progeny (Figures 2A and 2D; Figure S2B).

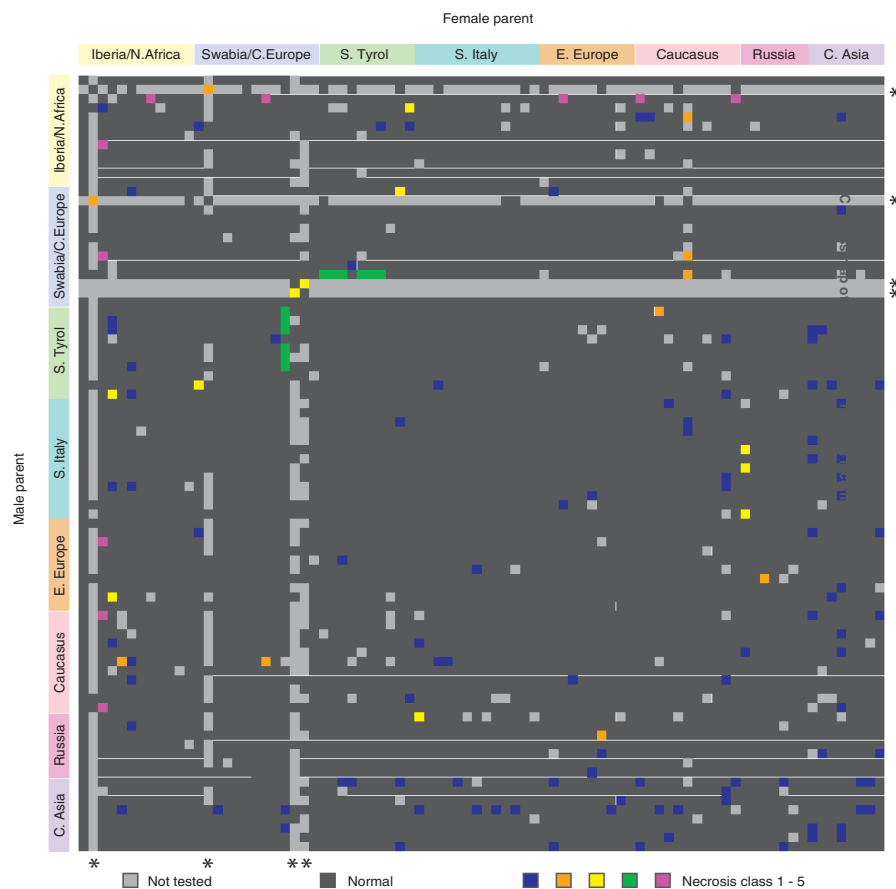


Figure 1. Crosses for Detection of Hybrid Incompatibilities

Accessions are color coded by region of origin. Accessions that are not part of the 80 accessions from Cao et al. (2011) are marked with asterisks. See also Figure S1 and Table S1.

ACCELERATED CELL DEATH 6 (ACD6), is reported elsewhere (Todesco et al., 2014).

We confirmed genes from the *RPP7* and *RPP4/5* NLR clusters as causal for *DM6* and *DM8* using artificial microRNAs (amiRNAs) (Table 1). *DM4* also overlapped the location of an *RPP* cluster, *RPP8* (Table S4). Four out of nine *DM* loci, accounting for ten of the causal alleles, thus map to highly variable *RPP* clusters (Figure 2I), which are the major sources for resistance to *Hpa* in *A. thaliana* (Nemri et al., 2010). In one *DM2* case, discussed in detail below, we have direct evidence for interactions between two different NLR genes. Three other interactions, *DM2-DM4*, *DM2-DM5*, and the *DM8* heterozygous incompatibility, also likely involve interactions between NLR genes, whereas the *DM6-DM7* interactions involve an NLR candidate and a complex non-NLR immune locus, *RPW8* (Xiao

Our analyses identified seven new hybrid necrosis loci, *DM3-DM9* (Figure 2; Table 1; Figure S2A), with final mapping intervals between 110 and 969 kb (Table S4). The *DM2* region was represented in multiple crosses: *DM3*, *DM4*, and *DM5* all interacted with *DM2* alleles from different strains, as do the previously identified *DM1* and *SRF3* loci (Alcázar et al., 2010; Bomblies et al., 2007). Thus, at least five out of the nine *A. thaliana* hybrid necrosis cases include *DM2* as one of the interactors. Two cases mapped to different pairs of *DM6* and *DM7* alleles, and two involved heterozygous effects at single loci, *DM8* and *DM9*.

We identified the *DM3* prolyl aminopeptidase (At3g61540) from Hh-0 as an interactor of *DM2h* from Bla-1 using transformation with genomic fragments and microRNA-induced gene silencing (MIGS) knockdown (Figure 2; Table 1; Figures S2B–S2E). In two crosses, KZ10 x Mrk-0 and Fei-0 x Lerik1-3, one of the causal loci, *DM7*, mapped to the *RESISTANCE TO POWDERY MILDEW 8 (RPW8)* region, which contains a variable tandem array of genes encoding coiled-coil proteins (Xiao et al., 2001). Transgenic experiments revealed that *RPW8.1^{KZ10}* was sufficient to induce necrosis in the Mrk-0 background (Figure 2; Table 1; Figures S2F and S2G). Despite similar genomic locations of the incompatibility genes, KZ10 is compatible with Lerik1-3, and Mrk-0 with Fei-0 (Table S1), indicating that these incompatibilities likely include different pairs of alleles at *DM6* and *DM7*.

We mapped the *DM9* locus in two crosses, Bla-1 x Hh-0 (this work) and Mir-0 x Se-0. A detailed analysis of the causal locus,

et al., 2001). Finally, the heterozygous interaction at *DM9* is caused by distinct alleles of another complex non-NLR immune locus, *ACD6* (Lu et al., 2003). Our systematic mapping efforts therefore indicate that NLR alleles along with other polymorphic immune genes located in tandem arrays are responsible for most intraspecific *F₁* incompatibilities in *A. thaliana*.

Multiple Incompatibilities due to the Complex *RPP1/DM2* NLR Locus

To understand how incompatibilities at *RPP* clusters evolve, we studied the *RPP1/DM2* locus from the accessions Uk-1 and Bla-1 in detail. We first assembled genomic sequences of the Uk-1 and Bla-1 *DM2* clusters from overlapping fosmid clones, and compared these with sequences from the Ler and Col-0 accessions and the related species *Arabidopsis lyrata*. In the *A. thaliana* reference strain Col-0, the *DM2* locus contains two *RPP1* paralogs that span 31 kb and that are part of a 180 kb *RPP1* supercluster. The *DM2* regions are much larger in Bla-1 and Uk-1, 119 kb and 128 kb (Figure 3A; Figure S3). Both accessions contain eight *RPP1* paralogs, similar to the 92 kb *DM2^{Ler}* cluster, which includes seven complete and at least one truncated *RPP1* homolog (Alcázar et al., 2009). Not a single *RPP1*-like gene is identical between accessions, consistent with the pattern of accession-specific incompatibilities (Figure 3B).

To test which of the *RPP1*-like genes in Uk-1 and Bla-1 are responsible for hybrid necrosis, we first knocked down individual

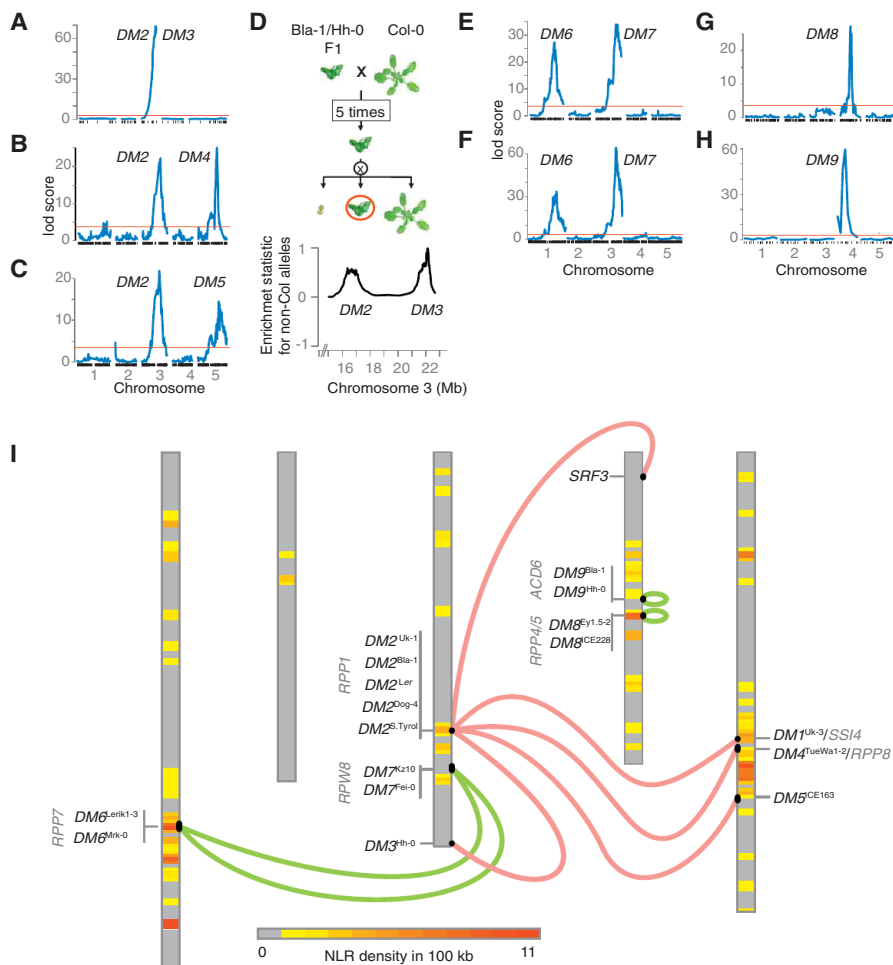


Figure 2. Linkage Mapping of Seven Hybrid Incompatibilities

(A–C and E–H) Hybrid necrosis QTL maps in (A) Bla-1/Hh-0 (leaf twisting), (B) TueWa1-2/ICE163, (C) Dog-4/ICE163, (E) Fei-0/Lerik1-3, (F) KZ10/Mrk-0, (G) Ey1.5-2/ICE228, and (H) Bla-1/Hh-0 (late-onset lesioning). Red lines mark the significance threshold ($p = 0.05$, 1,000 permutations); vertical marks along the x axes indicate marker positions.

(D) Scheme for Illumina sequencing of bulk segregants to delineate DM2 and DM3.

(I) Genomic location of DM alleles compared to NLR gene density (1 Mb windows) on the five chromosomes of the reference strain Col-0. SRF3, DM2^{Ler}, and DM1/SSI4 have been reported before (Alcázar et al., 2009, 2010; Bomblies et al., 2007). DM2/RPP1 interactions are in red; others are in green.

See also Figure S2 and Tables S2, S3, and S4.

that DM2 directly couples to downstream signaling events (Chung et al., 2011). These results demonstrate that incompatible pairs of DM proteins are sufficient to initiate cell-death signaling conserved between species.

Distinct Evolutionary Histories of Two Causal DM2 Alleles

Not all clades in a DM2 phylogeny (Figures S4H–S4J) include DM2 genes from all accessions. In addition, relationships among accessions within one clade often differ from those in another clade. Thus,

likely as a result of independent cycles of local gene duplication and loss along with illegitimate recombination and gene conversion (Table S5), DM2 clusters from different lineages show little conserved synteny, vary in size, and are poorly conserved outside NLR gene fragments (Figure 3A). This is consistent with patterns reported for major immune receptor gene clusters throughout flowering plants (Jacob et al., 2013).

Despite the overall similarity of DM2h/At3g44670 alleles among several accessions (Figures S4H–S4J), the LRR region, which is likely responsible for pathogen recognition, showed signs of diversifying selection (Figure 5A; Figures S5A and S5B; average substitutions per nonsynonymous site (K_a)/substitutions per synonymous site (K_s) = 4.2 for codons encoding the putative solvent-exposed residues). We thus hypothesized that rare allelic differences in a rapidly evolving portion of this gene gave rise to the incompatible behavior of DM2h^{Bla-1}. We localized residues responsible for incompatibility with DM3 using domain swaps. We first mapped the incompatible sequences to the C terminus of the DM2h^{Bla-1} protein, which includes the LRRs (Figure S5C). By engineering polymorphisms from DM2h^{Bla-1} that are rare in other accessions into the At3g44670^{Col-0} reference allele (Figures S5A and S5B), we identified two adjacent residues in the putative solvent-exposed site

DM2 genes with allele-specific amiRNAs (Figure S4A; Extended Experimental Procedures). We also introduced DM2 genomic clones into the incompatible parents Uk-3, which carries a DM2 interactor at DM1, and Hh-0, which carries a DM2 interactor at DM3 (Table 1). These experiments identified single genes in each accession, DM2d^{Uk-1} and DM2h^{Bla-1}, as necessary and sufficient for hybrid necrosis (Figure 4A; Figures S4B–S4G).

We asked whether autoimmunity depended on additional factors specific to the incompatible accessions. We first reconstituted the DM2^{Uk-1}-DM1^{Uk-3} and DM2^{Bla-1}-DM3^{Hh-0} interactions in the Col-0 background by crossing lines with the respective transgenes; in both cases, doubly transgenic lines were severely necrotic (Figure 4B). Next, we transiently expressed each pair in *Nicotiana benthamiana* leaves. We observed necrosis symptoms that mimicked the hypersensitive response (HR) seen upon recognition of a pathogen by a plant host when incompatible alleles of DM2 and DM1, or DM2 and DM3, were combined (Figure 4C). Importantly, DM2 genes closely related to either DM2d^{Uk-1} or DM2h^{Bla-1} did not trigger HR-like necrosis in *N. benthamiana* when combined with DM1^{Uk-3} or DM3^{Hh-0} (Figure 4C), confirming that HR is not simply induced by any combination of foreign NLR genes. Furthermore, enzymatic activity of DM2 was required for HR in this system (Figure 4C), indicating

Table 1. Candidate Genes for Hybrid Necrosis

Cross	Class ^a	Locus A	Evidence ^b	Locus B	Evidence ^b
Uk-1/Uk-3	3	<i>DM2 (RPP1)</i>	genomic/amiRNA	<i>DM1 (SSI4)</i>	genomic/amiRNA ^c
Bla-1/Hh-0 F ₃	2	<i>DM2 (RPP1)</i>	genomic/amiRNA	<i>DM3 (At3g61540)</i>	genomic/MIGS ^d
Bla-1/Hh-0 F ₂	2	<i>DM9 (ACD6)</i>	genomic/amiRNA ^e	<i>DM9 (ACD6)</i>	genomic/amiRNA ^e
TueWa1-2/ICE163	4	<i>DM2 (RPP1)</i>	map only	<i>DM4 (RPP8)</i>	map only
Dog-4/ICE163	2	<i>DM2 (RPP1)</i>	map only	<i>DM5</i>	map only
KZ10/Mrk-0	3	<i>DM6 (RPP7)</i>	map only	<i>DM7 (RPW8)</i>	genomic
Fei-0/Lerik1-3	2	<i>DM6 (RPP7)</i>	amiRNA	<i>DM7 (RPW8)</i>	map only
Ey1.5-2/ICE228	3	<i>DM8 (RPP4/5)</i>	amiRNA	<i>DM8 (RPP4/5)</i>	amiRNA

^aClasses 1–3 as described (Bomblies et al., 2007). Class 4 hybrids arrest as seedlings with only cotyledons and severe necrosis at 16°C; the seedlings were rescued by transfer to 28°C to set seeds.

^b“Genomic” refers to reconstitution of hybrid necrosis after transformation of genomic fragments into transgenic plants. See also Table S7 for details on the rescue experiments.

^cDescribed previously (Bomblies et al., 2007).

^dMIGS, miRNA-induced gene silencing (de Felippes et al., 2012).

^eDescribed in detail elsewhere (Todesco et al., 2014).

of LRR4 that can confer partial necrosis-inducing activity when combined with C-terminal sequences (Figure S5D). This result highlights the potential of natural NLR variants for the identification of residues that increase protein activity, which would inform efforts to engineer semisynthetic NLRs (Harris et al., 2013; Segretin et al., 2014).

RPP1-type proteins recognize and associate with proteins encoded by the *Hpa* ATR1 effector locus in an allele-specific manner (Krasileva et al., 2010). The hybrid necrosis-inducing residues in *DM2h*^{Bla-1} mapped near a modeled docking site of ATR1 onto RPP1-WsB (Steinbrener et al., 2012), suggesting that these variants have arisen from an arms race between an immune receptor and a pathogen ligand.

The topology of genes in the *DM2*^{Uk-1} cluster, as well as the phylogenetic relationships between *DM2* genes, suggested that *DM2d*^{Uk-1} arose by two within-cluster duplications and involved at least three gene conversion events (Figure 5B; Table S5). The two closest paralogs within the Uk-1 cluster are *DM2e* and *DM2g*, with *DM2e* having suffered mutations that truncate the encoded protein (Figure 5B; Figures S3 and S4H–S4J). Despite being within-cluster duplicates, *DM2d*^{Uk-1} and *DM2g*^{Uk-1} differ at many nonsynonymous sites, partly due to sequence exchanges with different paralogs (Figure 5A; Table S5). We visualized broader patterns of variation by using *DM2d*^{Uk-1} as a target for mapping of Illumina reads from 87 accessions. Accessions with very similar sequences across the entire gene were rare (Figure 5D). Moreover, similarity did not predict incompatibility: although ICE97 from southern Italy has a *DM2d* copy that is very similar to that of Uk-1, ICE97 was not incompatible with Uk-3 (Table S1). The rarity of *DM2d*^{Uk-1} is consistent with the hypothesis that *DM2d* is a rapidly evolving type I plant NLR gene, characterized by frequent sequence exchanges with other members of the same cluster (Kuang et al., 2004).

Unlike *DM2d*^{Uk-1}, *DM2h*^{Bla-1} shows a clear orthologous relationship with *DM2* genes in other accessions, *At3g44670*^{Col-0} and *At3g44670*^{Ler} (Figure 5C; Figures S4H–S4J), a pattern typical for type II plant NLR genes (Kuang et al., 2004). Alleles

at type II genes, which can be rare or common, diversify mostly by point mutations in the LRR region, rather than by sequence exchanges between paralogs. All three *DM2h*-type orthologs, *DM2h*^{Bla-1}, *At3g44670*^{Col-0}, and *At3g44670*^{Ler}, are located at the 3' end of the *DM2* cluster and mark the beginning of a syntenic region of at least 22 kb that is well conserved between Ler, Col-0, and Bla-1 but that is not found in Uk-1 or *A. lyrata* (Figures 3A and 6A). *DM2* hybrid necrosis alleles thus have arisen both as diversified orthologs and as paralogs within the tandemly arrayed gene cluster, accompanied by diversifying selection.

To assess the prevalence of haplotype sharing in this region, we asked whether the 3' syntenic region is present in 16 additional accessions with a *DM2h*-type gene (Figure 5D; Table S6). Reconstruction of phylogenetic relationships across an 8 kb region demonstrated that at least 12 of the *DM2h* carriers shared very similar sequences in this region (Figure 6B). Close relationships were not evident on the other side of the *DM2* cluster (Figure S6A), arguing against reduced haplotype diversity being simply a consequence of suppressed recombination, as reported for some NLR clusters (Chin et al., 2001). Such haplotype sharing among the 12 carriers, which extended over a region of 16 kb downstream of *DM2h/At3g44670* (Welch two-sample t test, $p < 0.0001$), was not observed next to two other NLR loci, the *RPM1* single-gene locus and the *RPP4/5* cluster (Figure 6C). We further confirmed haplotype sharing at *DM2* among the *DM2h* carriers using the F_{ST} statistic as a proxy for genetic differentiation (Figure S6B). Together with the observation that the 12 accessions are otherwise not particularly related either (Figure 6D), this suggests that the pattern of reduced haplotype diversity is not due to a recent population bottleneck.

Geographic Distribution of Hybrid Necrosis Risk Alleles

Two proteins that can produce hybrid necrosis in combination with *DM2* alleles from Uk-3 and Ler have been previously identified: the NLR protein DM1 from Uk-1 (Bomblies et al., 2007) and the kinase SRF3 from Central Asian accessions (Alcázar et al., 2010). In this study, we identified the prolyl aminopeptidase

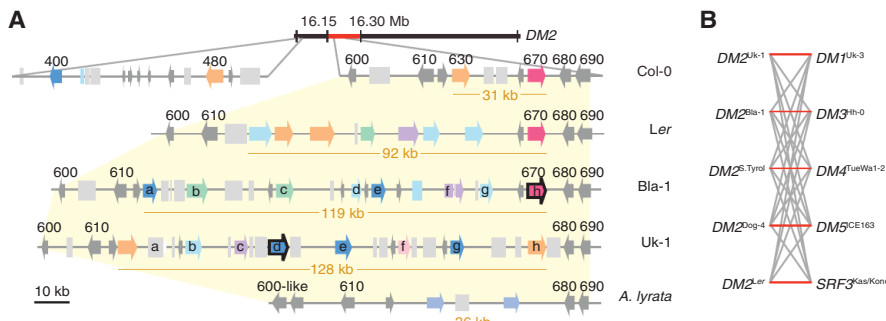


Figure 3. The *DM2* Cluster in *Arabidopsis*

(A) *DM2* clusters in four *A. thaliana* accessions and *A. lyrata* MN47. The mapping interval for *DM2* in Bla-1 is indicated in red. Genes are indicated with colored arrows, pseudogenes with colored boxes, and transposons with light gray boxes. Non-NLR genes are in dark gray. NLR genes are colored according to their phylogenetic history (see Figure S4), with unresolved relationships indicated in light blue. Numbers above the arrows indicate the last three digits of At3g44XXX and are given only when there are homologs in the reference genome. The two incompatibility genes are outlined in black. The *DM2*^{Ler} cluster was reannotated based on GenBank

accession number FJ446580.1. The sizes in kilobases refer to the core *DM2* clusters, defined as the coding regions of all NLRs (colored arrows) in a cluster. (B) Test crosses between *DM2* carriers and interacting allele carriers. Red and gray lines indicate incompatible and compatible interactions, respectively. See also Figure S3.

DM3 from Hh-0 as an interactor of *DM2h* from Bla-1 (Figure 6E; Figures S2B–S2E). In addition, the *DM2* cluster from Dog-4 interacts with an unknown gene at the *DM5* locus from ICE163, whereas the *DM2* clusters from several South Tyrolean accessions including ICE163 interact with *DM4* from TueWa1-2 (Figure 6E). We analyzed the genome-wide differentiation of *DM2* risk allele carriers among the 80 accessions that served as parents of many of our hybrid crosses plus other known carriers. As expected, accessions with different *DM2* alleles did not cluster with each other but rather with other accessions from the same geographical regions (Figure 6D), supporting independent origins of the different risk alleles. One of the *DM2* risk alleles was present in multiple strains from South Tyrol (Figure 6D; Table S1). This is similar to *SRF3*, for which incompatibility alleles are found throughout Central Asia (Alcázar et al., 2010), although overall population differentiation appears to be lower in Central Asia than in South Tyrol (Figures 6D and 6E) (Cao et al., 2011).

DISCUSSION

The extent to which nonadditive interactions between segregating alleles affect fitness-related traits in both outcrossing and selfing organisms is a central question in genetics (Corbett-Detig et al., 2013; Mackay, 2014; Phillips, 2008). We have used a systematically structured population of *F*₁ hybrids to determine the prevalence and causes of a common form of deleterious epistasis in plants, hybrid necrosis. A main finding is that interactions among immune genes are the most common cause of hybrid necrosis; this observation indicates that there are limits to the assembly of potentially favorable immune gene alleles in the same genetic background.

The crosses we investigated included parental pairs that were from the same location and sometimes closely related throughout the genome as well as geographically and genetically distant parents. We found that genome-wide genetic differentiation, which is correlated with geographic distance in *A. thaliana* (Cao et al., 2011), is not a good predictor for hybrid incompatibility. We interpret the occurrence of incompatibilities between accessions from the same geographic region as a sign that the incompatibilities on their own do not greatly affect the frequency of the individual causal alleles in the population. The genetic architectures we uncovered include interactions between one

locus and several other distinct loci (*DM2* with *DM1*, *DM3*, *DM4*, *DM5*, and *SRF3*), between different pairs of alleles at the same two loci (*DM6* with *DM7*), and between different alleles at the same locus (at *DM8* and *DM9*). This highlights that particular loci are disproportionately dangerous and can repeatedly cause independent deleterious epistatic interactions. Another important finding is that a large fraction of the hybrid necrosis alleles map to plant NLR immune receptor loci. Although we do not yet have proof that any of the specific alleles we have identified are beneficial in nature, the extreme variability of a subset of immune genes is in itself thought to be advantageous, particularly where resistance genes coevolve with extant pathogens (Holub, 2001; Michelmores and Meyers, 1998; Yang et al., 2013). Moreover, crop breeders have actively selected hybrid necrosis genes because they confer agronomically relevant resistances (Bomblied and Weigel, 2007; Krüger et al., 2002). This specific connection to the immune system sets our study apart from intraspecific studies in other systems where the evolutionary forces that give rise to deleterious epistasis remain largely unknown (Corbett-Detig et al., 2013).

It may not appear surprising that many hybrid necrosis genes encode NLR proteins, but two findings were unexpected: that *RPP* genes, which correspond to the major regions of the *A. thaliana* genome that encode resistance to the pathogen *Hpa* (Holub and Beynon, 1997; Nemri et al., 2010), are overrepresented, and that a single locus, *DM2/RPP1*, is involved in over half of all hybrid necrosis cases mapped to date. Among *RPP* genes, *RPP1* appears to be the most versatile locus, with alleles conferring resistance against many different *Hpa* genotypes and mediating different necrosis phenotypes (Holub and Beynon, 1997). That *DM2/RPP1* is at the same time a frequent trigger of autoimmunity underscores the potential dangers of a rapidly evolving immune system, both with respect to new mutations at this locus and upon outcrossing between accessions.

Because we found several interactions between different alleles at the same locus, our observations have implications not only for what has been called statistical epistasis, which is concerned with interactions between segregating polymorphisms, but also for functional epistasis, which refers to the allowed mutational paths of individual loci (Weinreich et al., 2005). Similar to *DM2/RPP1*, NLR genes are often arranged in tandem arrays.

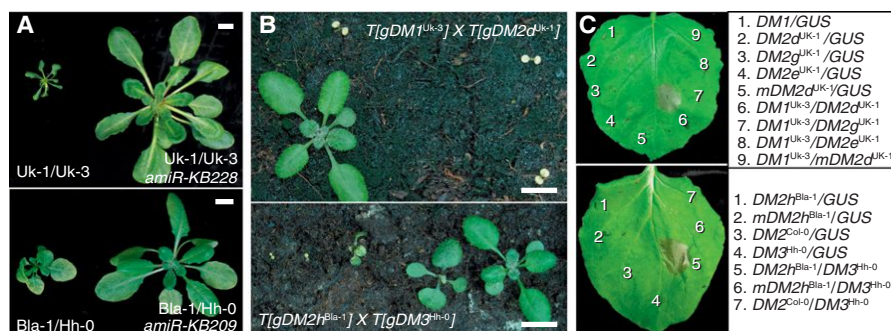


Figure 4. Identification of *DM2d^{Uk-1}* and *DM2h^{Bla-1}* as Hybrid Necrosis Genes

(A) *A. thaliana* F₁ hybrids and rescued siblings expressing amiRNAs at 16°C.

(B) Reconstitution of incompatibilities in *Col-0*, with transgenic F₁ hybrids at 16°C. Transgenic effects were often stronger than F₁ hybrid phenotypes.

(C) HR-like cell death in *N. benthamiana* induced by coexpression of *A. thaliana* DM proteins at 23°C. mDM2 indicates P loop mutant versions (GIGKTT to GIAATT), which should not be able to bind and hydrolyze ATP (Chung et al., 2011).

The scale bars represent 1 cm. See also Figure S4.

In a single tandem array, mutations could arise that cause deleterious interactions between proteins encoded by different members of such an array, but these would presumably be selected against, thereby limiting diversification within the array. In this context, it is of interest that the hybrid necrosis activity of the *DM2h^{Bla-1}* allele was apparently acquired stepwise, as deduced from our experiments with domain swaps.

Perhaps the most important conclusion from our findings is that autoimmunity in hybrids might limit the assembly of superior immune alleles into a single genotype, because of the interactions between NLRs and other loci involved in immunity. We note that the self-fertilizing mating system of *A. thaliana* is not a barrier to the rapid reassortment of genetic diversity. In the native range of the species, identical multilocus haplotypes are generally confined to individual small stands, and outcrossing rates in nature are sufficient to frequently generate new genetic combinations (Bomblies et al., 2010). It is noteworthy that we identified several accessions that carry multiple hybrid necrosis risk alleles. For example, the ICE163 accession carries both a *DM2* risk allele that is incompatible with *DM4* from TueWa1-2 and a risk allele at *DM5* that is incompatible with *DM2* from Dog-4. Similarly, hybrid necrosis alleles at both *DM2* and *DM9* are found in Bla-1, at both *DM4* and *DM7* in TueWa1-2, and at both *DM7* and an unmapped class 5 locus in TueScha-9 (Table S1). Multiple incompatibility risk alleles in the same genome would increase the chances of deleterious epistasis between immune genes upon crosses with other genotypes.

Although most hybrid necrosis alleles appear to be rare, we emphasize that we have limited our analyses to cases that are associated with strong morphological defects. These cases are almost certainly only the extremes of a wider range of interactions that lead to autonomous activation of the immune system. This argument follows from several observations. First, the F₁ hybrid necrosis cases display a range of severity, with some dying without flowering and others being able to set seeds as dwarves, and one of the *DM2* cases described in the literature is expressed only in the F₂ generation (Alcázar et al., 2009). Similarly, in our diallel among the 80 fully sequenced accessions, we have observed dozens of weakly necrotic F₁ cases, several of which showed stronger symptoms in the F₂ generation. Thus, it is likely that in addition to the interactions we have reported here there are others that cause milder immune phenotypes but still limit growth and development in a manner that is disad-

vantageous in the wild. Second, expression of hybrid necrosis symptoms can be modulated by genetic background (Alcázar et al., 2009), suggesting that more complex crossing strategies than the biparental design used here may reveal additional cases of hybrid necrosis.

Epistatic interactions between components of the immune system are likely to be relevant in other kingdoms as well. Such interactions have, for example, been observed in mammals, where there is evidence for positive selection acting on specific combinations of KIR-type receptors and major histocompatibility complex class I ligands (Single et al., 2007). Allelic variation at these loci is also responsible for autoimmune syndromes (Trowsdale and Knight, 2013). In *A. thaliana*, it seems perhaps unlikely that the interactions between the specific hybrid necrosis risk alleles we have described are beneficial, but it is conceivable that interaction between other alleles at these loci has positive effects on immune function. This can in principle be addressed using population genomic data, but because of the extreme variability of many of these loci, current whole-genome resequencing data sets are insufficient for asking directly whether specific alleles co-occur more often than expected by chance.

An important question for the future will be the biochemical nature of the interactions between hybrid necrosis risk alleles and how they differ from interactions between nonrisk alleles. *DM2* risk alleles trigger hybrid necrosis when combined with alleles at loci that encode a broad range of biochemical functions, including at least one NLR, *DM1*, consistent with other cases of NLR proteins that act in pairs (Eitas and Dangl, 2010). Two other natural *DM2* interactors encode a kinase (*SRF3*; Alcázar et al., 2010) and a prolyl aminopeptidase (*DM3*; this work). In addition, a chemically induced allele at a gene encoding a cysteine synthase can combine with a natural *DM2* allele to cause necrosis (Tahir et al., 2013). Outside of *A. thaliana*, hybrid necrosis alleles encode a cysteine protease (Krüger et al., 2002), a kinase (Yamamoto et al., 2010), and a subtilisin-like protease (Chen et al., 2014). Enzyme-encoding genes are clearly enriched, but in most cases we do not know yet how they modulate the activity of NLRs.

In conclusion, we propose that the study of hybrid necrosis can provide important insights into the role of epistatic interactions, particularly between immune genes, in the evolution of genotypes with multiple resistances to diverse pathogens. That hybrid necrosis alleles can increase functional disease

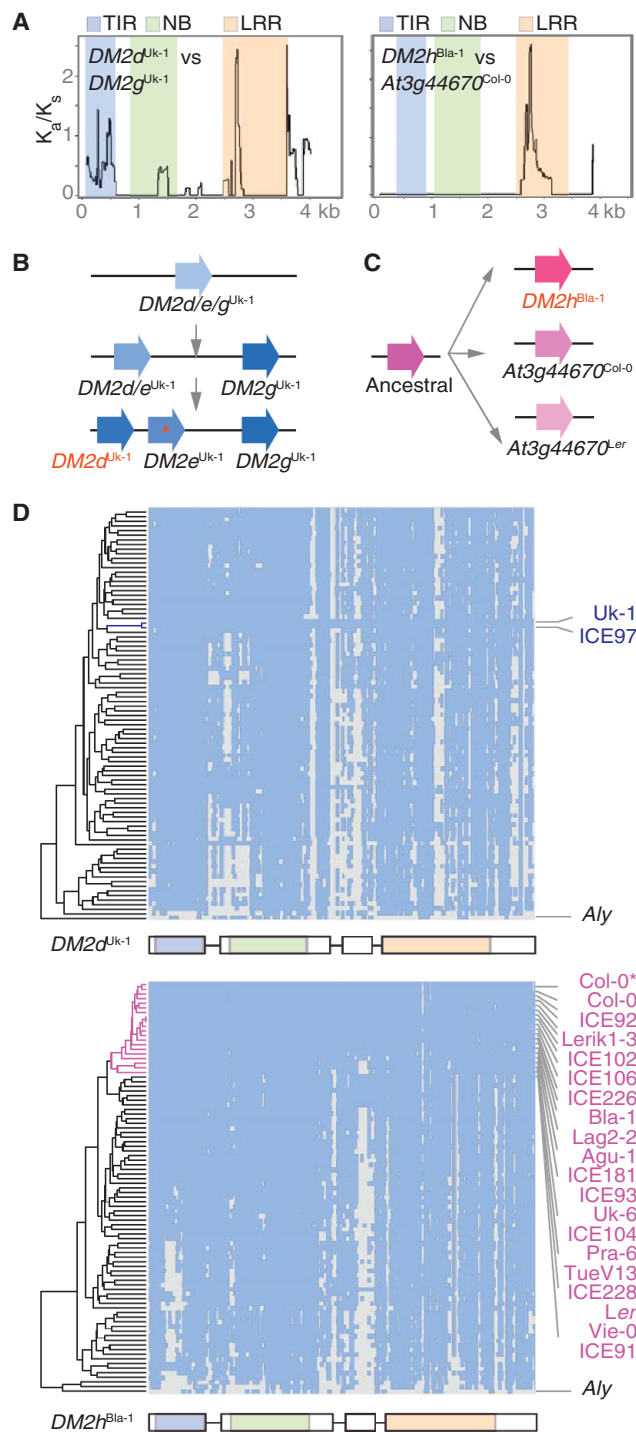


Figure 5. Origin and Variation of *DM2* Hybrid Necrosis Genes

(A) K_a/K_s ratios of closely related *DM2* genes (window length 150 bp, step size 9 bp). TIR, Toll/interleukin 1 receptor domain; NB, nucleotide-binding domain. (B) Most parsimonious path for evolution of *DM2d^{Uk-1}* paralogs. (C) Most parsimonious path for evolution of *DM2h^{Bla-1}* orthologs. (D) *DM2d^{Uk-1}*- and *DM2h^{Bla-1}*-type profiling using Illumina reads from accessions with one mismatch. Gray indicates uncovered regions. Accessions carrying a *DM2h* type are labeled in magenta. See also Figure S5 and Tables S5 and S6.

resistance in crop-breeding programs suggests that greater immune effectiveness may be tied to autoimmune risk. Understanding the relationship between effectiveness in pathogen recognition and autoimmunity will have applications in crop breeding, where it can help to guide strategies for maximizing disease resistance while minimizing yield penalties. Finally, it will be important to investigate whether the immune system of obligatory outbreeding species is more or less constrained than that of self-compatible species, and whether it therefore produces adverse effects in progeny less or more often.

EXPERIMENTAL PROCEDURES

Plant Material

F₁ hybrids were grown at 16°C under long days (16 hr of light). They were monitored for signs of autoimmunity-associated morphological defects for the first 12 weeks of growth. Afterward, plants were transferred to 23°C, and F₂ seeds were harvested from individual plants that did not carry the sterility-inducing *AP3* amiRNA transgene. Class 4 and 5 hybrids were additionally grown at 28°C.

Genotyping and QTL Analyses

A GBS approach was used for genotyping F₂ mapping populations, with PstI/MseI double-digested tags. Sequence tags and segregating SNPs for bulk segregants were generated either on a Illumina GAIIX or Illumina HiSeq 2000. Filtered markers were used for further QTL analyses (Table S3). Details on the identification of causal *DM* genes can be found in Table 1, Figures S2 and S4, and Table S7.

DM Sequence Analyses

A total of six fosmid clones each was Sanger shotgun sequenced to assemble the *DM2* locus in *Bla-1* and *Uk-1*. Illumina reads of 87 *A. thaliana* accessions (from Cao et al., 2011 plus *Col-0*, *Ler*, *Uk-1*, *Uk-3*, *Uk-6*, *Nc-1*, and *Bla-1*) and *A. lyrata* MN47 were trimmed to 36 bp in length and aligned to *DM2d^{Uk-1}* and *DM2h^{Bla-1}* open reading frames as reference using GenomeMapper (Schneeberger et al., 2009a). One mismatch and zero gaps were allowed. Matrices were generated by assigning a value of 1 to each position covered by at least one read, and a value of 0 to the remaining positions. Resulting profiles were clustered using complete linkage clustering with Euclidean distance.

Population Genetic Analyses

One hundred-kilobase regions upstream or downstream of three NLR loci were extracted from a genome variant matrix (Cao et al., 2011). Only positions with allele frequency above 0.1 were retained, and ten consecutive SNPs were binned for further calculations. The “compute” program, based on the “libsequence” C++ library (Thornton, 2003), was used to calculate haplotype diversity.

Oligonucleotides

Sequences of oligonucleotides used can be found in Table S8.

ACCESSION NUMBERS

Short reads of *Uk-3*, *Uk-6*, *KZ10*, and *Mrk-0* have been deposited in the European Nucleotide Archive under accession number ERP005469. Sequences of the *DM2* regions from *Bla-1* and *Uk-1*, *DM3* region from *Hh-0*, and *RPW8* region from *KZ10* have been deposited in GenBank under accession numbers KJ454428, KJ45449, KJ634210, and KJ634211, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, six figures, and eight tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2014.10.049>.

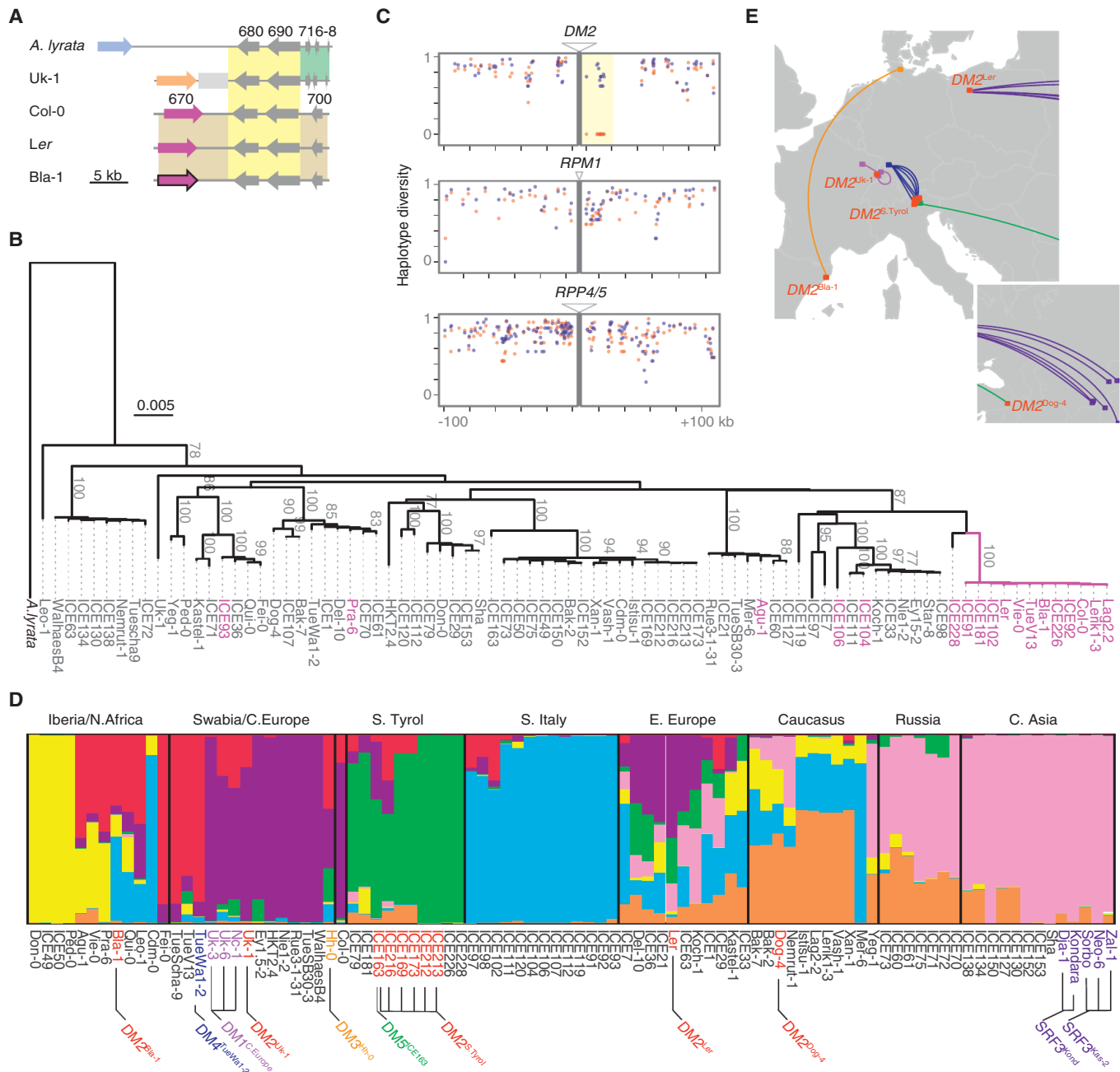


Figure 6. Haplotype Sharing around *DM2h* and Geographic Distribution of *DM2* Alleles

(A) Syntenic overview of the region beyond the *DM2* clusters in *A. thaliana* accessions and *A. lyrata* MN47. See the legend for Figure 3A for color and number coding.

(B) Phylogeny of 8 kb conserved sequences spanning At3g44680 and At3g44690 (yellow shading in A). *DM2h*-type carriers are in magenta and noncarriers are in gray. Bootstrap values over 70% are indicated.

(C) Haplotype diversity based on groups of ten adjacent SNPs in the regions flanking NLR loci *DM2*, *RPM1*, and *RPP4/5*. Twelve *DM2h*-type carriers are in red and 12 noncarriers are in blue.

(D) STRUCTURE analysis (version 2.3.4; Pritchard et al., 2000) ($k = 7$) of hybrid necrosis risk allele carriers together with a selection of global accessions. At the bottom, accessions carrying different *DM2* hybrid necrosis alleles are in red, and those carrying *DM2*-interacting alleles are in other colors. These colors are unrelated to the ones used to identify membership in STRUCTURE clusters at the top.

(E) Carriers of *DM2* hybrid necrosis alleles are in red; carriers of interacting alleles are in other colors. *DM2*-*SRF3* interactions are from Alcázar et al. (2010). See also Figure S6.

AUTHOR CONTRIBUTIONS

E.C., K.B., and D.W. conceived and designed the experiments; E.C., K.B., S.-T.K., M.Z., C.M.-P., H.T., S.L., M.D., A.H.-M., and C.L. performed the experiments; E.C., K.B., S.-T.K., D.K., M.Z., S.O., R.A.E.L., B.A.R., and G.R. analyzed the data; and E.C., K.B., and D.W. wrote the paper, with contributions from all authors.

ACKNOWLEDGMENTS

We thank Rubén Alcázar, Jane Parker, and Maarten Koornneef for information regarding *RPP1*, Jesse Poland for GBS advice, Frank Wellmer for *AP3* amiRNA, Eui-Hwan Chung and Jeffery Dangl for pointers on cell-death assays, William Ho, Paula Sancha-Vilchez, and Josip Perkovic for technical support, and Jeffery Dangl, Ya-Long Guo, Daniel Koenig, George Wang, and Jun Cao for discussions. We especially thank the anonymous reviewers, who greatly helped us with the evolutionary framing of our work. This work was supported by an NIH Ruth Kirschstein National Research Service Award (K.B.), Human Frontier Science Program Long-Term Fellowship (R.A.E.L.), Alexander von Humboldt Foundation Fellowship (B.A.R.), Human Frontier Science Program Organization Grant (RGP 57/2007), and a Gottfried Wilhelm Leibniz Award of the Deutsche Forschungsgemeinschaft and the Max Planck Society (D.W.).

Received: April 15, 2014

Revised: July 22, 2014

Accepted: October 7, 2014

Published: November 20, 2014

REFERENCES

- Alcázar, R., García, A.V., Parker, J.E., and Reymond, M. (2009). Incremental steps toward incompatibility revealed by *Arabidopsis* epistatic interactions modulating salicylic acid pathway activation. *Proc. Natl. Acad. Sci. USA* 106, 334–339.
- Alcázar, R., García, A.V., Kronholm, I., de Meaux, J., Koornneef, M., Parker, J.E., and Reymond, M. (2010). Natural variation at Strubbelig Receptor Kinase 3 drives immune-triggered incompatibilities between *Arabidopsis thaliana* accessions. *Nat. Genet.* 42, 1135–1139.
- Allen, R.L., Bittner-Eddy, P.D., Grenville-Briggs, L.J., Meitz, J.C., Rehmany, A.P., Rose, L.E., and Beynon, J.L. (2004). Host-parasite coevolutionary conflict between *Arabidopsis* and downy mildew. *Science* 306, 1957–1960.
- Bakker, E.G., Toomajian, C., Kreitman, M., and Bergelson, J. (2006). A genome-wide survey of *R* gene polymorphisms in *Arabidopsis*. *Plant Cell* 18, 1803–1818.
- Bateson, W. (1909). Heredity and variation in modern lights. In Darwin and Modern Science, A.C. Seward, ed. (Cambridge, UK: Cambridge University Press), pp. 85–101.
- Bikard, D., Patel, D., Le Metté, C., Giorgi, V., Camilleri, C., Bennett, M.J., and Loudet, O. (2009). Divergent evolution of duplicate genes leads to genetic incompatibilities within *A. thaliana*. *Science* 323, 623–626.
- Bombles, K., and Weigel, D. (2007). Hybrid necrosis: autoimmunity as a potential gene-flow barrier in plant species. *Nat. Rev. Genet.* 8, 382–393.
- Bombles, K., Lempe, J., Eppele, P., Warthmann, N., Lanz, C., Dangl, J.L., and Weigel, D. (2007). Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. *PLoS Biol.* 5, e236.
- Bombles, K., Yant, L., Laitinen, R.A., Kim, S.T., Hollister, J.D., Warthmann, N., Fitz, J., and Weigel, D. (2010). Local-scale patterns of genetic variability, outcrossing, and spatial structure in natural stands of *Arabidopsis thaliana*. *PLoS Genet.* 6, e1000890.
- Cao, J., Schneeberger, K., Ossowski, S., Günther, T., Bender, S., Fitz, J., Koenig, D., Lanz, C., Stegle, O., Lippert, C., et al. (2011). Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nat. Genet.* 43, 956–963.
- Chen, C., Chen, H., Lin, Y.S., Shen, J.B., Shan, J.X., Qi, P., Shi, M., Zhu, M.Z., Huang, X.H., Feng, Q., et al. (2014). A two-locus interaction causes interspecific hybrid weakness in rice. *Nat. Commun.* 5, 3357.
- Chin, D.B., Arroyo-Garcia, R., Ochoa, O.E., Kesseli, R.V., Lavelle, D.O., and Micheltore, R.W. (2001). Recombination and spontaneous mutation at the major cluster of resistance genes in lettuce (*Lactuca sativa*). *Genetics* 157, 831–849.
- Chung, E.H., da Cunha, L., Wu, A.J., Gao, Z., Cherkis, K., Afzal, A.J., Mackey, D., and Dangl, J.L. (2011). Specific threonine phosphorylation of a host target by two unrelated type III effectors activates a host innate immune receptor in plants. *Cell Host Microbe* 9, 125–136.
- Collier, S.M., and Moffett, P. (2009). NB-LRRs work a “bait and switch” on pathogens. *Trends Plant Sci.* 14, 521–529.
- Corbett-Detig, R.B., Zhou, J., Clark, A.G., Hartl, D.L., and Ayroles, J.F. (2013). Genetic incompatibilities are widespread within species. *Nature* 504, 135–137.
- Coyne, J.A., and Orr, H.A. (2004). Speciation (Sunderland, MA: Sinauer Associates).
- Crespi, B., and Nosil, P. (2013). Conflictual speciation: species formation via genomic conflict. *Trends Ecol. Evol.* 28, 48–57.
- Cutter, A.D. (2012). The polymorphic prelude to Bateson-Dobzhansky-Muller incompatibilities. *Trends Ecol. Evol.* 27, 209–218.
- de Felippes, F.F., Wang, J.W., and Weigel, D. (2012). MIGS: miRNA-induced gene silencing. *Plant J.* 70, 541–547.
- Dobzhansky, T. (1937). Genetics and the Origin of Species (New York: Columbia University Press).
- Eitas, T.K., and Dangl, J.L. (2010). NB-LRR proteins: pairs, pieces, perception, partners, and pathways. *Curr. Opin. Plant Biol.* 13, 472–477.
- Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., and Mitchell, S.E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6, e19379.
- Harris, C.J., Slootweg, E.J., Goverse, A., and Baulcombe, D.C. (2013). Step-wise artificial evolution of a plant disease resistance gene. *Proc. Natl. Acad. Sci. USA* 110, 21189–21194.
- Holub, E.B. (2001). The arms race is ancient history in *Arabidopsis*, the wild-flower. *Nat. Rev. Genet.* 2, 516–527.
- Holub, E.B., and Beynon, J.L. (1997). Symbiology of mouse-ear cress (*Arabidopsis thaliana*) and oomycetes. *Adv. Bot. Res.* 24, 227–273.
- Hou, J., Friedrich, A., de Montigny, J., and Schacherer, J. (2014). Chromosomal rearrangements as a major mechanism in the onset of reproductive isolation in *Saccharomyces cerevisiae*. *Curr. Biol.* 24, 1153–1159.
- Jacob, F., Vernaldi, S., and Maekawa, T. (2013). Evolution and conservation of plant NLR functions. *Front. Immunol.* 4, 297.
- Jeuken, M.J., Zhang, N.W., McHale, L.K., Pelgrom, K., den Boer, E., Lindhout, P., Micheltore, R.W., Visser, R.G., and Niks, R.E. (2009). *Rin4* causes hybrid necrosis and race-specific resistance in an interspecific lettuce hybrid. *Plant Cell* 21, 3368–3378.
- Krasileva, K.V., Dahlbeck, D., and Staskawicz, B.J. (2010). Activation of an *Arabidopsis* resistance protein is specified by the in planta association of its leucine-rich repeat domain with the cognate oomycete effector. *Plant Cell* 22, 2444–2458.
- Krüger, J., Thomas, C.M., Golstein, C., Dixon, M.S., Smoker, M., Tang, S., Mulder, L., and Jones, J.D. (2002). A tomato cysteine protease required for Cf-2-dependent disease resistance and suppression of autonecrosis. *Science* 296, 744–747.
- Kuang, H., Woo, S.S., Meyers, B.C., Nevo, E., and Micheltore, R.W. (2004). Multiple genetic processes result in heterogeneous rates of evolution within the major cluster disease resistance genes in lettuce. *Plant Cell* 16, 2870–2894.
- Lachance, J., and True, J.R. (2010). X-autosome incompatibilities in *Drosophila melanogaster*: tests of Haldane's rule and geographic patterns within species. *Evolution* 64, 3035–3046.

- Lu, H., Rate, D.N., Song, J.T., and Greenberg, J.T. (2003). ACD6, a novel ankyrin protein, is a regulator and an effector of salicylic acid signaling in the *Arabidopsis* defense response. *Plant Cell* 15, 2408–2420.
- Mackay, T.F. (2014). Epistasis and quantitative traits: using model organisms to study gene-gene interactions. *Nat. Rev. Genet.* 15, 22–33.
- Maheshwari, S., and Barbash, D.A. (2011). The genetics of hybrid incompatibilities. *Annu. Rev. Genet.* 45, 331–355.
- Michelmore, R.W., and Meyers, B.C. (1998). Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Res.* 8, 1113–1130.
- Muller, H.J. (1942). Isolating mechanisms, evolution and temperature. *Biol. Symp.* 6, 71–125.
- Nemri, A., Atwell, S., Tarone, A.M., Huang, Y.S., Zhao, K., Studholme, D.J., Nordborg, M., and Jones, J.D. (2010). Genome-wide survey of *Arabidopsis* natural variation in downy mildew resistance using combined association and linkage mapping. *Proc. Natl. Acad. Sci. USA* 107, 10302–10307.
- Phillips, P.C. (2008). Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* 9, 855–867.
- Phillips, P.C., and Johnson, N.A. (1998). The population genetics of synthetic lethals. *Genetics* 150, 449–458.
- Poland, J.A., Brown, P.J., Sorrells, M.E., and Jannink, J.L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE* 7, e32253.
- Presgraves, D.C. (2010). The molecular evolutionary basis of species formation. *Nat. Rev. Genet.* 11, 175–180.
- Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Quintana-Murci, L., and Clark, A.G. (2013). Population genetic tools for dissecting innate immunity in humans. *Nat. Rev. Immunol.* 13, 280–293.
- Rieseberg, L.H., and Blackman, B.K. (2010). Speciation genes in plants. *Ann. Bot. (Lond.)* 106, 439–455.
- Sackton, T.B., Lazzaro, B.P., Schlenke, T.A., Evans, J.D., Hultmark, D., and Clark, A.G. (2007). Dynamic evolution of the innate immune system in *Drosophila*. *Nat. Genet.* 39, 1461–1468.
- Schneeberger, K., Hagmann, J., Ossowski, S., Warthmann, N., Giesing, S., Kohlbacher, O., and Weigel, D. (2009a). Simultaneous alignment of short reads against multiple genomes. *Genome Biol.* 10, R98.
- Segretin, M.E., Pais, M., Franceschetti, M., Chaparro-Garcia, A., Bos, J.I., Banfield, M.J., and Kamoun, S. (2014). Single amino acid mutations in the potato immune receptor R3a expand response to *Phytophthora* effectors. *Mol. Plant Microbe Interact.* 27, 624–637.
- Seidel, H.S., Rockman, M.V., and Kruglyak, L. (2008). Widespread genetic incompatibility in *C. elegans* maintained by balancing selection. *Science* 319, 589–594.
- Single, R.M., Martin, M.P., Gao, X., Meyer, D., Yeager, M., Kidd, J.R., Kidd, K.K., and Carrington, M. (2007). Global diversity and evidence for coevolution of *KIR* and *HLA*. *Nat. Genet.* 39, 1114–1119.
- Steinbrenner, A.D., Goritschnig, S., Krasileva, K.V., Schreiber, K.J., and Staskawicz, B.J. (2012). Effector recognition and activation of the *Arabidopsis thaliana* NLR innate immune receptors. *Cold Spring Harb. Symp. Quant. Biol.* 77, 249–257.
- Tahir, J., Watanabe, M., Jing, H.C., Hunter, D.A., Tohge, T., Nunes-Nesi, A., Brotman, Y., Fernie, A.R., Hoefgen, R., and Dijkwel, P.P. (2013). Activation of *R*-mediated innate immunity and disease susceptibility is affected by mutations in a cytosolic *O*-acetylserine (thiol) lyase in *Arabidopsis*. *Plant J.* 73, 118–130.
- Thornton, K. (2003). libsequence: a C++ class library for evolutionary genetic analysis. *Bioinformatics* 19, 2325–2327.
- Todesco, M., Balasubramanian, S., Hu, T.T., Traw, M.B., Horton, M., Eppe, P., Kuhns, C., Sureshkumar, S., Schwartz, C., Lanz, C., et al. (2010). Natural allelic variation underlying a major fitness trade-off in *Arabidopsis thaliana*. *Nature* 465, 632–636.
- Todesco, M., Kim, S.T., Chae, E., Bomblies, K., Zaidem, M., Smith, L.M., Weigel, D., and Laitinen, R.A. (2014). Activation of the *Arabidopsis thaliana* immune system by combinations of common *ACD6* alleles. *PLoS Genet.* 10, e1004459.
- Trowsdale, J., and Knight, J.C. (2013). Major histocompatibility complex genomics and human disease. *Annu. Rev. Genomics Hum. Genet.* 14, 301–323.
- Vilches, C., and Parham, P. (2002). KIR: diverse, rapidly evolving receptors of innate and adaptive immunity. *Annu. Rev. Immunol.* 20, 217–251.
- Weinreich, D.M., Watson, R.A., and Chao, L. (2005). Perspective: sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* 59, 1165–1174.
- Wuest, S.E., O'Maoileidigh, D.S., Rae, L., Kwasniewska, K., Raganelli, A., Hanczaryk, K., Lohan, A.J., Loftus, B., Graciet, E., and Wellmer, F. (2012). Molecular basis for the specification of floral organs by APETALA3 and PISTILLATA. *Proc. Natl. Acad. Sci. USA* 109, 13452–13457.
- Xiao, S., Ellwood, S., Calis, O., Patrick, E., Li, T., Coleman, M., and Turner, J.G. (2001). Broad-spectrum mildew resistance in *Arabidopsis thaliana* mediated by *RPW8*. *Science* 291, 118–120.
- Yamamoto, E., Takashi, T., Morinaka, Y., Lin, S., Wu, J., Matsumoto, T., Kitano, H., Matsuoka, M., and Ashikari, M. (2010). Gain of deleterious function causes an autoimmune response and Bateson-Dobzhansky-Muller incompatibility in rice. *Mol. Genet. Genomics* 283, 305–315.
- Yang, S., Li, J., Zhang, X., Zhang, Q., Huang, J., Chen, J.Q., Hartl, D.L., and Tian, D. (2013). Rapidly evolving *R* genes in diverse grass species confer resistance to rice blast disease. *Proc. Natl. Acad. Sci. USA* 110, 18572–18577.

Paternal Diet Defines Offspring Chromatin State and Intergenerational Obesity

Anita Öst,^{1,6,8,*} Adelheid Lempradl,^{1,8} Eduard Casas,^{2,3} Melanie Weigert,¹ Theodor Tiko,¹ Merdin Deniz,¹ Lorena Pantano,² Ulrike Boenisch,¹ Pavel M. Itskov,⁷ Marlon Stoeckius,⁴ Marius Ruf,¹ Nikolaus Rajewsky,⁴ Gunter Reuter,⁵ Nicola Iovino,¹ Carlos Ribeiro,⁷ Mattias Alenius,⁶ Steffen Heyne,¹ Tanya Vavouri,^{2,3} and J. Andrew Pospisilik^{1,*}

¹Max Planck Institute of Immunobiology and Epigenetics, Stuebeweg 51, 79108 Freiburg, Germany

²Institut de Medicina Predictiva i Personalitzada del Càncer, Can Ruti Campus, Ctra. de Can Ruti, Camí de les Escoles s/n, 08916 Badalona, Barcelona, Spain

³Josep Carreras Leukaemia Research Institute (IJC), ICO-Hospital Germans Trias i Pujol, 08916 Badalona, Barcelona, Spain

⁴Systems Biology of Gene Regulatory Elements, Max-Delbrück-Center for Molecular Medicine, Robert-Rössle-Strasse 10, 13125 Berlin, Germany

⁵Institute of Biology, Developmental Genetics, Martin Luther University Halle, 06120 Halle, Germany

⁶Department of Clinical and Experimental Medicine, Linköping University, 58183 Linköping, Sweden

⁷Champalimaud Neuroscience Programme, Champalimaud Centre for the Unknown, 1400-038 Lisbon, Portugal

⁸Co-first author

*Correspondence: anita.ost@liu.se (A.Ö.), pospisilik@ie-freiburg.mpg.de (J.A.P.)

<http://dx.doi.org/10.1016/j.cell.2014.11.005>

SUMMARY

The global rise in obesity has revitalized a search for genetic and epigenetic factors underlying the disease. We present a *Drosophila* model of paternal-diet-induced intergenerational metabolic reprogramming (IGMR) and identify genes required for its encoding in offspring. Intriguingly, we find that as little as 2 days of dietary intervention in fathers elicits obesity in offspring. Paternal sugar acts as a physiological suppressor of variegation, desilencing chromatin-state-defined domains in both mature sperm and in offspring embryos. We identify requirements for H3K9/K27me3-dependent reprogramming of metabolic genes in two distinct germline and zygotic windows. Critically, we find evidence that a similar system may regulate obesity susceptibility and phenotype variation in mice and humans. The findings provide insight into the mechanisms underlying intergenerational metabolic reprogramming and carry profound implications for our understanding of phenotypic variation and evolution.

INTRODUCTION

Global incidence of obesity is approaching 1 billion humans. Though poorly understood, parental and fetal nutritional states have been shown to generate reproducible offspring phenotypes, including obesity. Studies in multiple model organisms have been used to examine intergenerational metabolic effects (Braunschweig et al., 2012; Ozanne et al., 1999; Morgan et al., 1999; Buescher et al., 2013; Rechavi et al., 2014). Maternal and paternal induction of intergenerational responses have been reported, and a variety of macronutrient and timing interventions have been

used, including short- and long-term fasting (Anderson et al., 2006), calorie restriction (Blondeau et al., 2002; Jimenez-Chillaron et al., 2009), and modulation of dietary protein (Ozanne et al., 1999), fat (Gniuli et al., 2008; Dunn and Bale, 2009), and methyl-donor content (Wolff et al., 1998; Waterland et al., 2006; reviewed in Daxinger and Whitelaw, 2012; Patti, 2013). Of note, although not understood, divergent physiological extremes can prompt similar offspring phenotypes, so called “U-shaped” responses.

Intergenerational effects transmitted via the male germline have received recent attention. Because father-to-offspring transmission excludes difficult to control oocyte and gestational effects, mechanistic dissections are simplified. Studies have demonstrated paternal transmission of tumor susceptibility (Anway et al., 2005; Xing et al., 2007), of heat-shock-induced epigenetic memory (Seong et al., 2011), of olfaction-dependent behavioral and neural phenotypes (Dias and Ressler, 2014), and of metabolic control (Anderson et al., 2006; Fullston et al., 2013; Carone et al., 2010; Ng et al., 2010; reviewed in Rando, 2012).

Mechanistically, imprinting, altered DNA methylation, histone modifications, and noncoding RNA transcripts have been implicated in inter/transgenerational phenotype transmission. Adiposity of genetically identical agouti mouse siblings correlates with IAP DNA methylation (Morgan et al., 1999); DNA methylation correlates with endocrine disruptor and nutrient induced inter/transgenerational phenotypes (Anway et al., 2005; Carone et al., 2010; Radford et al., 2014), and there is evidence of RNA-dependent transmission (Gapp et al., 2014; Rechavi et al., 2014; Kiani et al., 2013; Rassoulzadegan et al., 2006). In *C. elegans* and *Drosophila*, research has focused on small non-coding RNAs and chromatin organization (Seong et al., 2011; Shirayama et al., 2012; Lee et al., 2012; Greer et al., 2011; Ashe et al., 2012). Despite these advances, however, our understanding of the initiation, transmission, and stabilization of trans/intergenerational phenotypes remains largely a black box.

Here, we present a *Drosophila* model of paternal intergenerational metabolic reprogramming (IGMR) and identify

germline and zygotic gene networks that are necessary for its manifestation. Mechanistically, paternal sugar modifies offspring chromatin state and transcription in a *Polycomb*-, *E(z)*-, *SetDB1*-, *Su(var)3-9*-, and *HP1*-sensitive manner. Intriguingly, these changes are forecast in the sperm. Data from highly defined human and mouse obesity cohorts suggest that these processes are conserved. These data provide evidence for a conserved chromatin-state-encoded program that defines phenotypic variation and thus carry profound implications for our understanding of phenotypic diversity and evolution.

RESULTS

A *Drosophila* Model of Intergenerational Metabolic State Control

We sought to understand whether normal fluctuations in diet might impact next-generation phenotypes. We chose to focus on the male germline and, for simplicity, on male progeny.

To minimize genetic variation, we performed ten generations of single-fly, brother-sister inbreeding of our population inbred *w¹¹¹⁸* *Drosophila melanogaster* strain. To identify an optimal dietary intervention for P0 fathers, we challenged 4- to 5-day-old males with progressively increasing dietary sugar and protein and assessed whole-fly fat storage after 2 days. Whereas dietary protein showed minimal effects (Figure 1A; horizontal axis), dietary sugar evoked a 3-fold increase in whole-fly triglyceride storage (Figure 1A; vertical axis). Of note, the sugar concentrations used approximate natural food sources (ripened banana ~300 g/l). These responses agreed with published data (Skorupa et al., 2008) and highlighted the rapid metabolic regulatory potential of dietary sugar in the fly.

To test for intergenerational effects, we repeated the experiment, this time varying only sugar, and mated the males to standardized *w¹¹¹⁸* female virgins (Figure 1B). After 2 hr of mating, females were left to lay eggs for 10 hr, removed, and the F1 offspring were left to develop unimpeded. Importantly, ancestral (more than ten generations), parental, and F1 generations were highly controlled with respect to male:female mating ratio, larval density, diet, and environmental conditions. One week after eclosion, adult male offspring were weighed and sacrificed, and triglyceride levels were measured in whole-fly lysates. Interestingly, although the paternal intervention showed no effect on F1 kept on normal food (Figure 1C, top, open circles), adult progeny fed an obesogenic high-sugar diet exhibited a U-shaped obesity response (Figure 1C, top closed circles) with low- and high-sugar sired individuals showing exaggerated triglycerides (Figure 1C and 1D). This phenotype was significant by both ANOVA and comparison of linear versus polynomial regressions. Thus, paternal sugar outside of the physiological optimum alters metabolic control in the F1.

The IGMR phenotype comprised two features. F1 offspring body weight increased with paternal sugar (Figure 1C, middle), and weight-normalized triglyceride levels increased toward both paternal extremes (Figure 1C and 1D). Metabolic phenotyping revealed that obesity-susceptible IGMR progeny exhibited increased adipose area (Figure 1E and 1F) and lipid droplet size (Figure 1E and Figure S1A available online). Measures of feeding behavior showed a tendency toward increased food

intake (Figure 1G). Together with increased starvation sensitivity (Figure S1D), unaltered activity and CO₂ production measures (Figures S1B and S1C), as well as unaltered trehalose and glucose levels (Figures S1E and S1F), these findings suggest that the observed excess triglyceride reserves resulted from poor lipid store mobilization and possibly hyperphagia. Importantly, we found no evidence of altered eclosion timing (Figure 1H) or wing size (Figure 1I, top), or in F1 offspring number (Figure 1I, middle) or male:female ratio per brood (Figure 1I, bottom). Thus, acute paternal dietary sugar reprograms offspring metabolism, leaving growth and development intact. Our data conclusively provide evidence that acute paternal diet reprograms offspring metabolism in *Drosophila*.

Paternal IGMR Is Rapid, Stable, and Stress Sensitive

The short 2 day intervention implied that mature *Drosophila* sperm are capable of continuously transmitting environmental cues to their offspring. To understand the minimum dietary intervention required to elicit paternal IGMR, *w¹¹¹⁸* males were subject to dietary interventions lasting 1, 2, 5, or 7 days prior to mating (Figure 2A). Measurements of adiposity indicated a maximal F1 phenotypic response within just 2 days of paternal challenge (Figures 2B and 2C), suggesting that paternal IGMR might be detectable within a single day. Female flies store sperm upon mating, allowing them to fertilize eggs for days or even weeks after a single insemination event. We asked how stable the IGMR phenotype would be with sperm storage. Males were subject to a 2 day paternal dietary intervention and mated, and the recipient females were allowed to lay three consecutive batches of fertilized eggs over 60 hr. Importantly, offspring of all three consecutive batches exhibited clear U-shaped obesity phenotypes (Figures 2D and 2E), indicating that IGMR is stable with sperm storage. Further, heat shock is known to modulate epigenetically controlled phenotypes, even across generations (Seong et al., 2011). Interestingly, a 1 hr heat shock of diet-treated fathers immediately before mating completely abrogated the paternal IGMR obesity response (Figures 2F and 2G). This is consistent with the requirement for highly controlled environmental conditions (sound, odor, vibration) when using this model. Finally, we found no evidence that the IGMR phenotype is transmitted to subsequent generations (Figure 2H). Thus, paternal IGMR is rapid, stable, and acutely stress sensitive.

Paternal Sugar Alters Offspring Heterochromatin

Paternal IGMR appeared phenotypically “silent” through the complexities of development. We therefore hypothesized that the phenotype was encoded in chromatin. Position-effect variegation (PEV) is a genetic phenomenon that has been used as a quantitative readout of locus-specific chromatin state silencing in vivo. The most common PEV reporters in *Drosophila* reflect chromatin desilencing through increased expression of a red-eye-pigment-coding reporter gene. Screening a library of PEV strains, Phalke and colleagues recently defined at least five functionally distinct chromatin silencing subtypes in the living fly (Phalke et al., 2009). Using identical or comparable PEV lines, we tested whether paternal dietary sugar could alter offspring eye color and thus stably alter chromatin state in F1 (Figures 3A–3E). We observed no overt effect of paternal diet on offspring

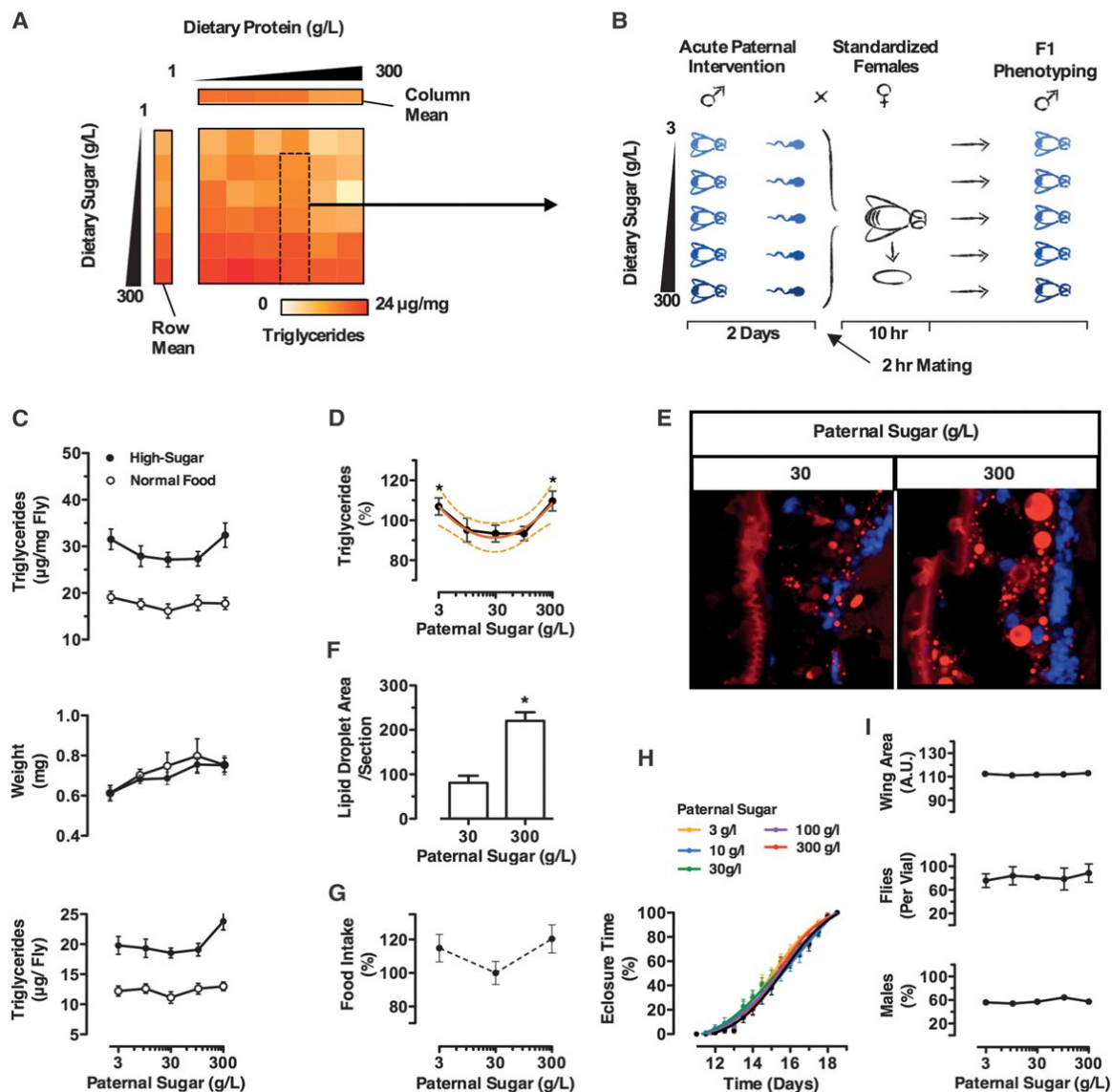


Figure 1. A Fly Model of Paternally Induced Obesity

(A) Triglyceride levels in founder males after 2 days of diet intervention with increasing sugar (sucrose) and protein (soy).
 (B) Schematic of the IGMR experimental design.
 (C) Triglycerides (top), body weight (middle), and weight normalized triglycerides (bottom) of F1 males raised on normal (open circles) and on high-sugar food (closed circles).
 (D) Binomial regression of F1 male weight normalized triglycerides (95% CI, $p < 0.01$).
 (E) Representative section of anterior fly fat body stained with oil red O and DAPI.
 (F) Lipid droplet area/section.
 (G) Food intake of F1 males by CAFE assay.
 (H) Eclosion timing of F1 offspring as percent of total.
 (I) (Top) Relative wing area of F1 males. (Middle) Total number of eclosed offspring per vial. (Bottom) Number of male offspring per vial (% of total flies). Results are mean \pm SEM ($p < 0.05$) of $n = 3$ –8 experiments each with multiple replicates. See also Figure S1.

PEV in four of the lines tested (A_{480} ; Figures 3A–3D), including reporters for telomeric (ChrX; *HA-1902*) (Figure 3A), retro-transposon-type (Chr3R; *HA-1992*) (Figure 3B), pericentric (Chr4;39c-12) (Figure 3C), and repeat-associated chromatin (Chr2:3;92E) (Figure 3D). Notably, all four lines generated U-shaped paternal IGMR obesity (Figure 3F). Thus, IGMR occurs

on independent genetic backgrounds and leaves *HA-1902*-, *HA-1992*-, 39c-12-, and 92E-type chromatin largely unaltered.

Intriguingly, when testing *w^{m4h}*, a reporter for peri-centric heterochromatin on ChrX, we observed a reproducible U-shaped intergenerational eye color phenotype (Figures 3E and 3G). In support of a mechanistic link between the IGMR obesity and

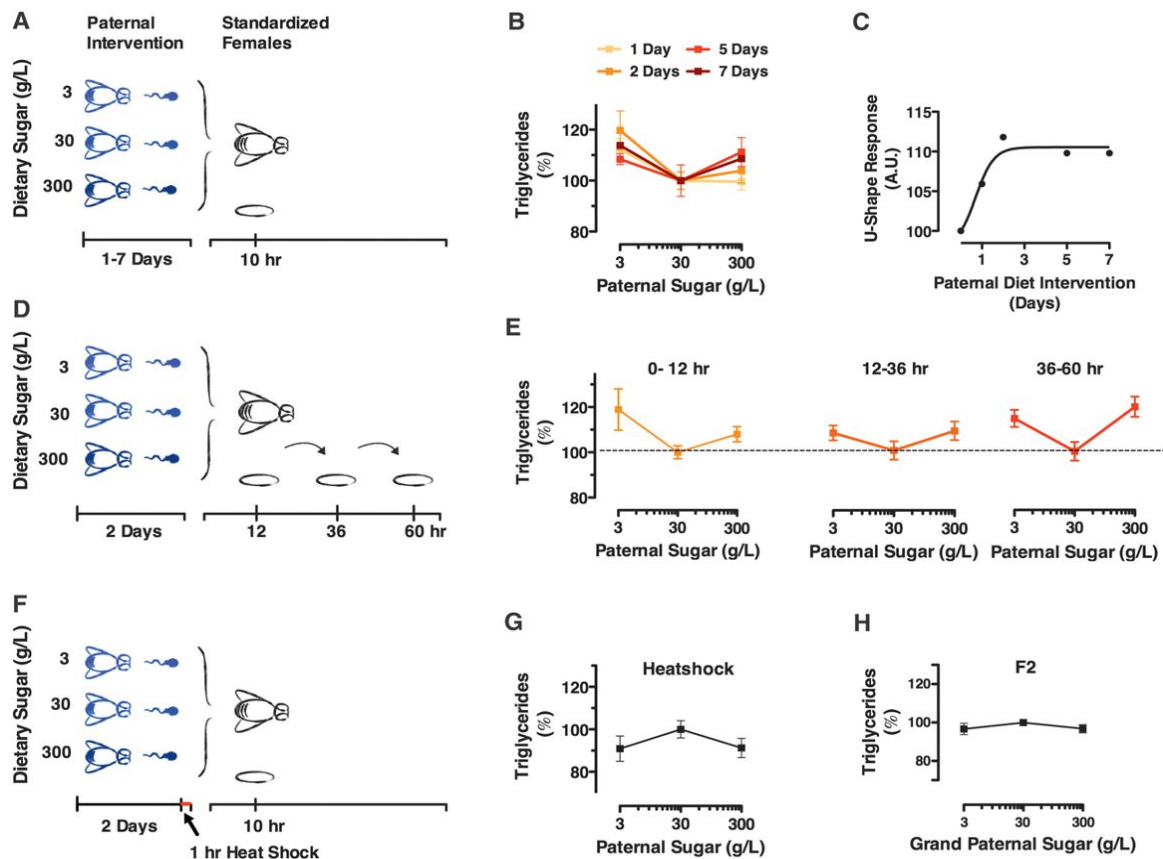


Figure 2. Acute Paternal Nutritional State Is Reflected in Offspring Obesity

(A and B) (A) Schematic and (B) offspring adiposity from tests of progressively increasing paternal dietary intervention.

(C) Mean change in offspring adiposity (Δ triglycerides/weight) for low- and high-sugar-sired adult males relative to medium sugar for each time point. Least square curve fitting (slope = -1 , $R^2 = 0.95$).

(D and E) (D) Schematic and (E) F1 male adult adiposity from tests of consecutive offspring cohorts from the same mating event. After mating, females were kept on standard food, and three consecutive batches of embryos were collected and assessed at adulthood.

(F and G) (F) Schematic and (G) offspring adiposity for tests of stress sensitivity. A 1 hr 37 degree heat-shock was applied to males just before mating.

(H) F2 adult male adiposity. F1 males were kept on standard food prior to mating.

Results are mean \pm SEM of $n = 3-8$ experiments each with multiple replicates.

PEV results, triglyceride accumulation and eye color correlated positively; redder-eyed flies were more obese (Figure 3H). No correlation was observed in the remaining four strains (data not shown). These data show that acute paternal diet targets select chromatin subtypes in offspring.

High Paternal Sugar Controls Heterochromatin-Embedded Gene Expression

At this point, we focused on medium- versus high-sugar IGMR and tested whether IGMR affected all or only select individuals in the population. Measuring pigment from single w^{m4h} fly heads as a direct readout of the IGMR response, we observed that paternal IGMR red-shifted the entire distribution (Figures 4A and 4B). Thus, high paternal sugar induces w^{m4h} desilencing population-wide, indicating that each paternal gamete carries an equivalent intergenerational signal.

Next, we performed rRNA-depleted RNA-sequencing of hand-picked stage 17 embryo F1 offspring from medium- and high-

sugar challenged fathers (Pearson corr. = 0.97, ~ 15 million reads/sample; Figure 4C). In support of a selective chromatin state desilencing mechanism, gene expression broadly increased, with many more up- than downregulated transcripts. Sixty-eight protein-coding genes were significantly upregulated in high-sugar sired embryos (mean Δ FPKM = 54.9) and only ten downregulated (mean Δ FPKM = -7.0 ; Figure 4D and Table S1). Of note, upregulated transcripts tended to be genes highly expressed during late embryo and early larval stages, including 27 (40%) related to biogenesis of the sugar-based cuticle. Of the remaining 42 genes, 30 were of unknown function, 5 had peptidase activity, and interestingly, 4 were metabolic genes, including fatty acyl-CoA reductase and fatty acid elongase.

Analysis using gene set enrichment analysis (GSEA) revealed two clearly upregulated clusters containing *chitin* and *cuticle constituent* and *mitochondrial and primary energy metabolism* pathways (Figures 4E and S2A and Table S2). Included and consistent with the heightened adiposity of IGMR, pathways

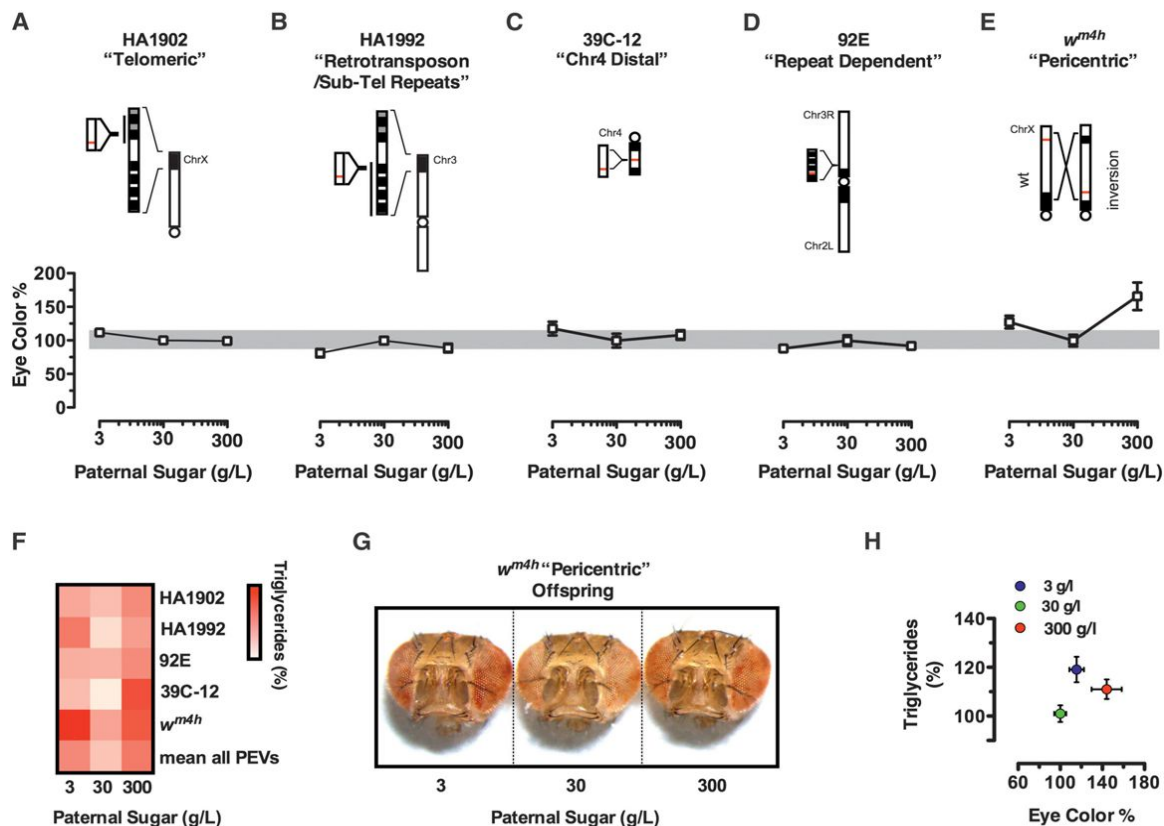


Figure 3. High Paternal Sugar Is a Physiological *Su(var)*

(A–E) (Top) Schematics of (A) *pP{RS5}5-HA-1902*, (B) *pP{RS5}5-HA-1992*, (C) *39C-12*, (D) *T(2;3)V21ePlacW* (92E), (E) *In(1)w^{m4h}* PEV reporters. (Bottom) Eye pigment absorption (A_{480}) from heads of adult males, normalized to offspring of paternal 30 g/L sugar.

(F) Heatmap of paternal IGMR offspring from the PEV lines in (A–E). Triglyceride/weight normalized to paternal 30 g/L sugar.

(G) Representative median eye-colored *w^{m4h}* fly heads from 3, 30, and 300 g/L sugar-sired offspring.

(H) Correlation of eye color (heads) and offspring triglyceride/weight (carcass rest) of *w^{m4h}* flies measured in the same individuals.

Results are mean \pm SEM of $n = 3$ –8 experiments each with multiple replicates.

for *lipid particle*, the electron transport chain complexes I, IV, and V, *glycolysis*, *TCA cycle*, and *fatty acid metabolism* were all upregulated. These changes are consistent with energetics of enhanced lipid storage (Figure S2B). Three downregulated clusters were also detected, including *cell cycle and mitosis*, *body patterning*, and intriguingly, a cluster of *chromatin regulation* pathways. Consistent with sensitivity of the pericentric *w^{m4h}* reporter to IGMR chromosome, "centromeric region" was ranked second in the chromatin cluster and "chromatin silencing" ranked third. Examination of genes annotated as PEV suppressing, also known as *Su(var)*'s, revealed a concerted $\sim 10\%$ – 20% downregulation, including members of most well-documented silencing pathways (Figure S2C and Tables S2 and S3). Thus, the IGMR embryo is characterized by gene expression favoring primary energy metabolism over chromatin control.

We next compared our data with chromatin mapping data sets from the community. Filion et al. used Dam-ID to annotate five major chromatin types, three repressive (black, blue, and green) and two active (red and yellow) (Filion et al., 2010). When intersecting our IGMR embryo data with their chromatin state maps, strong enrichment was observed in high-sugar sired embryos for genes

embedded in "black" lamin/H1-associated heterochromatin and "blue" polycomb-associated chromatin, and relative depletion was observed for those annotated as "yellow," or housekeeping-type chromatin (Figures 4F and 4G). These findings were verified using rank-order (Figure 4F) and differential expression analyses (Figure 4G). No global effect was observed on "red" or "green" chromatin embedded genes. Consistent with these global indications of chromatin state dependency, the 68 significantly upregulated genes were almost exclusively found in "black" or "blue" chromatin while the 10 significantly downregulated transcripts were randomly distributed (Figure 4H). These data identify high paternal sugar as a chromatin-state-selective physiological *Su(var)* and identify IGMR as chromatin state dependent.

Polycomb and Core Heterochromatin Machinery Mediate Paternal IGMR

To genetically validate chromatin state regulation as a mechanistic underpinning of our model, we began systematically testing IGMR potential in mutants known to modify *w^{m4h}* variegation. We started with *Su(var)3-9⁰⁶*, a homozygous dominant suppressor allele of the H3K9 histone methyltransferase *Su(var)3-9*.

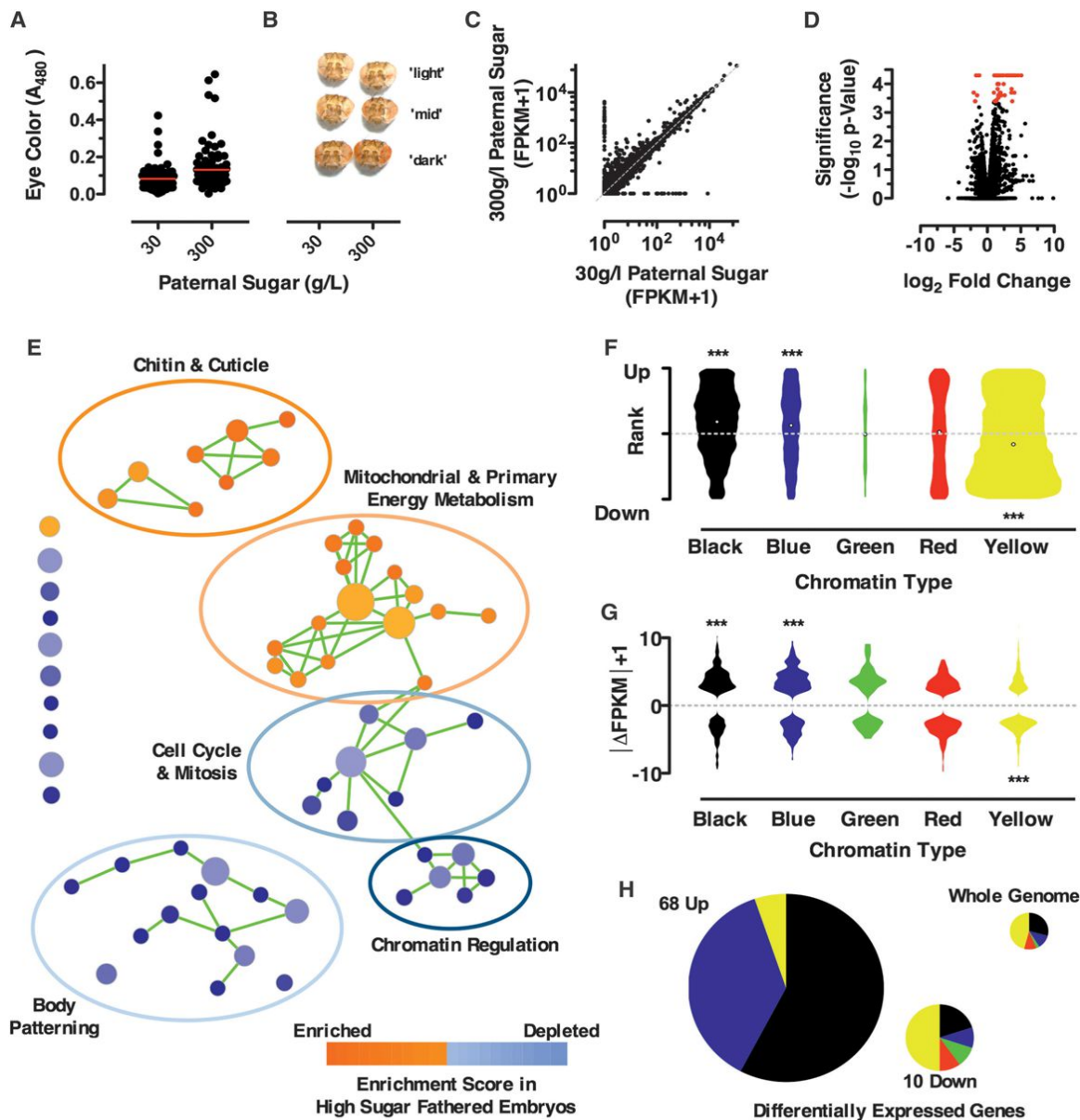


Figure 4. Paternal IGMR Alters Select Chromatin States in Offspring

(A) Interindividual variation of eye color of w^{1118} flies from fathers fed high (300 g/L) or medium (30 g/L) sugar food.

(B) Representative heads from w^{1118} offspring. Shown are heads representative of light, medium, and dark red eyes of each respective cohort.

(C–H) RNA-sequencing results of medium (30 g/L) and high-sugar (300 g/L) sired stage 17 embryos. (C) FPKM plot. (D) Volcano plot.

(E) Cytoscape enrichment map (p cutoff: 0.005, FDR Q-value cutoff: 0.025, overlap cutoff: 0.2) of gene set enrichment analysis (GSEA). (Orange) Gene sets enriched; (blue) gene sets depleted, in high-sugar IGMR. Color intensity reflects degree of enrichment. Major clusters are circled.

(F and G) (F) Rank and (G) absolute IGMR expression changes. Genes are allocated to one of five chromatin states (colors) according to their TSS (Filion et al., 2010). Plotted are (F) ranks for all genes and (G) absolute expression changes of the top 1,000 IGMR up and downregulated genes.

(H) Chromatin color annotation of all significantly up- and downregulated IGMR genes.

See also Figure S2.

Medium- and high-sugar-challenged *Su(var)3-9⁰⁶* fathers were mated with standardized w^{1118} females, and the resulting heterozygote offspring were monitored for adiposity (Figure 5A). Whereas w^{1118} animals reproducibly exhibited a ~10%–15% increase in adiposity upon high-sugar IGMR, F1 adult male offspring of *Su(var)3-9⁰⁶* fathers showed no intergenerational

obesity response (Figure 5A). This provides genetic evidence that *Su(var)3-9* is required for IGMR.

We also tested a second H3K9 methyltransferase, *SetDB1*. As heterozygotes, *SetDB1¹⁴⁷³* fathers gave both wild-type and mutant offspring. Intriguingly, both mutant (Figure 5B, red) and wild-type *SetDB1¹⁴⁷³* fathered offspring (Figure 5B, black)

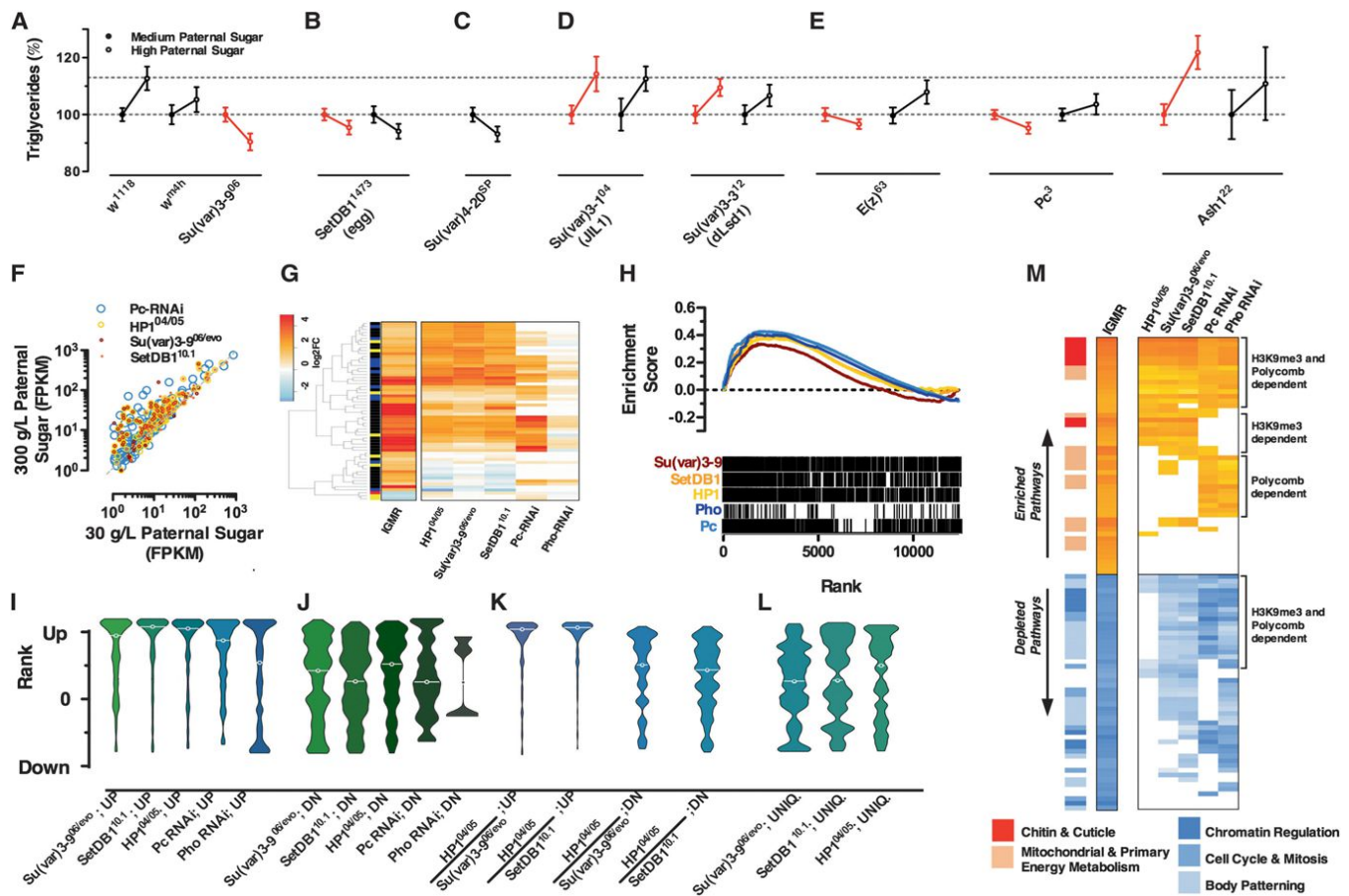


Figure 5. A *Su(var)/PcG* Axis Essential for Paternal IGMR

(A–E) Adiposity of offspring (triglycerides/weight) of mutant fathers challenged with medium (30 g/l; closed circles) or high sugar (300 g/l; open circles). Gray dashed line indicates normal *w¹¹¹⁸* IGMR response. IGMR adiposity responses are shown for offspring of (A) *w¹¹¹⁸*, *w^{m4h}*, and *Su(var)3-9⁰⁶*, (B) *SetDB1¹⁴⁷³*, (C) *Su(var)4-20^{SP}*, (D) *Su(var)3-1⁰⁴*, *Su(var)3-3¹²*, (E) *E(z)⁶³*, *Pc³*, and *Ash1²²* mutant (red) and wild-type (black) offspring. Results are mean \pm SEM of $n = 3-8$ experiments each with multiple replicates.

(F) FPKM values of RNaseq data from medium and high-sugar-fathered embryos. 200 most upregulated genes from HP1 (open yellow), *Su(var)3-9* (closed cayenne), and *SetDB1* (closed orange) mutant first-instar larvae from Lundberg et al. (2013) and *Pc*-RNAi (open blue) experiment from Goodliffe et al. (2007).

(G) Heatmap of expression changes of significantly changed genes in our paternal IGMR offspring embryo data set and in the Lundberg et al. *HP1*, *Su(var)3-9*, and *SetDB1* mutants and the Goodliffe et al. (2007) *Pc*-, *Pho*-RNAi data sets.

(H) Enrichment plot for gene sets upregulated in *HP1^{04/05}* (yellow), *Su(var)3-9^{06/06}* (cayenne), and *SetDB1^{10.1}* (orange) mutants and *Pc*- (light blue) and *Pho*-RNAi (dark blue) in our stage 17 paternal IGMR offspring embryos.

(I–L) Violin plots of expression change distributions relative to all genes of stage 17 paternal IGMR offspring embryos for gene sets from *HP1^{04/05}*, *Su(var)3-9^{06/06}*, and *SetDB1^{10.1}* mutants (Lundberg et al.) and from *Pc* and *Pho*-RNAi embryos (Goodliffe et al.). IGMR relative rank is plotted for all available of the (I) 200 genes most upregulated; and (J) 200 genes most downregulated by each mutant / RNAi line; (K) intersects of the 200 most up- or downregulated genes of the indicated pairs of mutants; and (L) genes in the 200 most upregulated gene sets unique to each respective mutant.

(M) Heatmap comparison of GSEA results from mutant and IGMR data. Plotted are the 50 most up- and downregulated pathways from paternal IGMR and respective scores from the mutant data sets. Colored bars left of the heatmap indicate clusters in Figure 4E.

completely failed to mount an IGMR obesity response (Figure 5B). *Drosophila* sperm develop as a syncytium, and therefore both mutant and wild-type sperm in such a cross will share a *SetDB1¹⁴⁷³* mutant cytosolic compartment for most of their development. These findings therefore indicate that *SetDB1* in the male germline is necessary for proper IGMR.

H4K20me3 deposition follows H3K9me3 in the establishment of heterochromatin (Schotta et al., 2004). We therefore also tested *Su(var)4-20^{SP}*, a mutant for the H4K20 methyltransferase *Su(var)4-20*. As *Su(var)4-20^{SP}* is on Chr X, all male offspring from

our crosses are wild-type. Again though, wild-type offspring will reflect the mutant heterozygosity of spermatogenesis. *Su(var)4-20^{SP}* fathers failed to transmit paternal IGMR to the F1 (Figure 5C). Thus, uncompromised expression of *Su(var)3-9*, *SetDB1*, and *Su(var)4-20* are absolutely required for IGMR. Of note, not all *w^{m4h}* suppressor alleles were IGMR incompetent. *Su(var)3-1⁰⁴* and *Su(var)3-3¹²*, also known as Jil1 kinase and dLSD1, respectively, generated completely normal IGMR obesity responses (Figure 5D), thus indicating that IGMR is not directly linked to the *w^{m4h}* insertion locus itself. These findings

identify one of the first gene networks known to be absolutely required for proper intergenerational metabolic reprogramming.

Given the observed derepression in blue embedded genes (polycomb-associated; Figure 4), we tested IGMR potential in polycomb and trithorax group mutants. We found that, although *Ash1*²² mutants were fully IGMR competent, *Enhancer of zeste*, *E(z)*⁶³ and *Polycomb*, *Pc*³ mutant males completely failed to elicit a response in the next generation (Figure 5E). Thus, polycomb- and H3K9me3-centric chromatin regulators are absolutely required for paternal diet-induced intergenerational obesity.

The IGMR Program Is Chromatin Encoded

To corroborate these findings, we compared our embryonic IGMR RNA-seq data with profiles from H3K9me3- and polycomb-insufficient mutants. We examined profiles from *Su(var)3-9*^{06/evo}, *SetDB1*^{10.1}, and *HP1*^{04/05} mutant first-instar larvae (Lundberg et al., 2013) and *Pc*- and *Pho*-RNAi knockdown embryos (Goodliffe et al., 2007). Intriguingly, ~70% overlap was observed between our significantly dysregulated IGMR genes and those responsive to H3K9-centric or polycomb insufficiency (Figures 5F and 5G). The converse was equally true; each of the top 200 *Su(var)3-9*^{06/evo}, *SetDB1*^{10.1}, *HP1*^{04/05}, *Pho*-RNAi, and *Pc*-RNAi dysregulated gene sets showed strong enrichment in our high-sugar-sired F1 embryos (Figure 5H). Subgrouping confirmed specificity of these signals. First, transcripts upregulated by *Su(var)3-9*^{06/evo}, *SetDB1*^{10.1}, *HP1*^{04/05}, *Pho*, and *Pc* insufficiency (likely direct targets) showed clear coordinate increases in expression (Figure 5I) compared to apparently randomly distributed signals for transcripts downregulated by mutation (Figure 5J). Transcripts upregulated by both *HP1*^{04/05} and either *Su(var)3-9*^{06/evo} or *SetDB1*^{10.1} (Figure 5K) showed much stronger signatures than transcripts significantly regulated by any one *Su(var)* mutant alone (Figure 5L). Thus, paternal IGMR mimics H3K9me3- and polycomb-dependent transcriptional dysregulation.

To test whether these signatures might directly contribute to metabolic reprogramming, we performed GSEA analysis of the Lundberg et al. (2013) and Goodliffe et al. (2007) data sets. Coordinate overlapping enrichment signatures were observed for key pathways of all five major IGMR clusters (Figure 5M and Table S2), including most chromatin and primary energy modules. Of note, the most significantly enriched pathways in our data set were those regulated by both silencing systems together (Figure 5M). Thus, IGMR is characterized by H3K9me3-/PcG-dependent dysregulation.

Sperm and Zygote Chromatin Plasticity Define IGMR

To gain further insight into IGMR transmission, we performed RNA sequencing from manually dissected and purified mature sperm of high- and medium-sugar-fed *w*¹¹¹⁸ males (Table S4). Intriguingly, we again observed clear evidence of (1) broad transcriptional derepression in sperm of high-sugar-fed males (Figures 6A and 6B), (2) selective upregulation of black chromatin-embedded genes (Figure 6C), and (3) upregulation of *Su(var)3-9*^{06/evo}-sensitive genes (Figure 6D). These data indicate that transcriptional dysregulation in mature IGMR sperm is also chromatin state defined. In contrast to the embryo data, blue and yellow embedded genes appeared largely unaffected in

the sperm transcriptome. Thus, chromatin-dependent signatures of IGMR are forecast in the P0 paternal germline.

Dysregulation of black embedded genes in both the sperm and zygote suggested potentially overlapping mechanisms for generation of the intergenerational signal in the germline and for hardwiring the IGMR phenotype in the offspring. To probe this idea genetically, we compared the effect of maternal versus paternal mutant allele contribution on IGMR. As described above, offspring of *Su(var)* and *Polycomb* mutant fathers were incapable of mounting an IGMR response (Figures 5A–5E and 6E, top row). In crosses in which *Su(var)3-9*^{06/evo}, *SetDB1*^{10.1}, or *Su(var)4-20*^{SP} mutations were contributed by the oocyte, IGMR-competent wild-type sperm were no longer able to evoke an intergenerational response (Figure 6E, bottom row). *Su(var)3-104* and *Su(var)3-3*¹² mutants, unremarkable in the male germline, completely abrogated the response when contributed maternally (Figure 6E, bottom row). In contrast, oocytes contributing *Pc*³ and *E(z)*⁶³ mutations, whose constitutive heterochromatin would not be predicted to be directly perturbed, mounted completely normal IGMR responses. Collectively, these data support a model in which IGMR results from and requires a permissive range of heterochromatin plasticity in the zygote.

To validate the idea, we intersected our embryo RNA-seq data with modENCODE H3K9me3 and H3K27me3 ChIP-seq profiles from same-stage embryos (16–20 hr) and from those isolated one time point earlier in development (12–16 hr), enabling us to gauge the dynamics of H3K9me3/K27me3 gain and loss (Nègre et al., 2011). We made several observations. First, IGMR-dysregulated genes represented a class undergoing highly dynamic H3K9 and H3K27 trimethylation (Figure 6F). This was true for our significantly changed IGMR genes, as well as the leading edge H3K9me3- and polycomb-dependent IGMR gene sets from Figure 5I (Figure 6F and data not shown). The bodies of these genes in particular were unmarked in 12–16 hr embryos and exhibit strong H3K9me3 and H3K27me3 just 4 hr later. Importantly, we observed the same signature when analyzing leading-edge genes of metabolic pathways upregulated in our obese IGMR phenotype (Figure S3A). Thus, genes undergoing highly dynamic H3K9me3- and H3K27me3-dependent silencing are specifically targeted for IGMR derepression.

Because repressive marks correlate with the higher-order chromatin structure and *cis*-regulatory domain organization, we also examined our gene sets in the context of insulator occupancy (Nègre et al., 2010). Analysis revealed that all three IGMR-dysregulated gene sets were on average far from class I (*CTCF*, *CP190*, and *BEAF*-associated) and were somewhat closer to class II (*SuHw*-associated) insulators (Figure 6G). These signatures were specific when compared to similarly expressed genes or to the entire transcriptome (Figure 6H, left). Intriguingly, the same signature was again evident in our most up- and downregulated sperm transcripts (Figure 6H, right). Thus, IGMR impacts spatially and chromatin-context-defined transcriptional units in fathers and in offspring.

Collectively, our data suggest that IGMR results from global alterations in chromatin state integrity within a permissive window, where obesity susceptibility results from reduced stage-specific epigenetic regulation of H3K27me3- and H3K9me3-defined domains. Our observations of *w*^{m4h} eye color desilencing (Figures 3

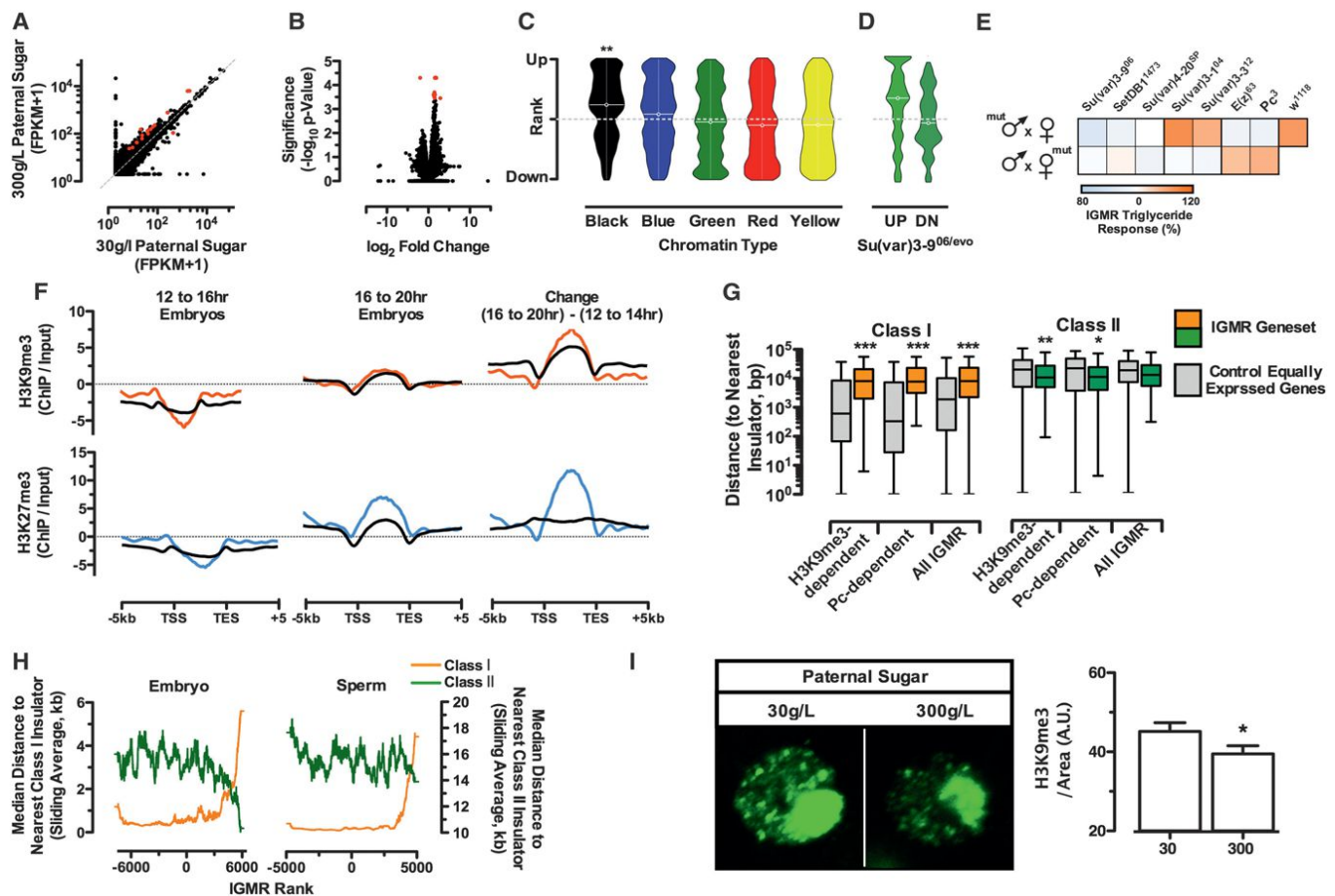


Figure 6. IGMR Signatures Are Forecast in the P0 Germline, and IGMR Changed Genes Show K9/K27me3 Dynamic Context

(A–D) RNA-sequencing results of sperm from medium- (30 g/L) and high-sugar (300 g/L) fed fathers; significantly changed genes are depicted in red. (A) FPKM plot. (B) Volcano plot. (C and D) IGMR expression changes in sperm of high- relative to medium-sugar-fed fathers (C) for the five chromatin colors according to Filion et al. (2010) and (D) for 200 most up- or downregulated genes from *Su(var)3-906/ev0* mutants from Lundberg et al (2013).

(E) Relative adiposity of male offspring (triglycerides/weight) (top row) from crosses of mutant fathers challenged with medium- or high-sugar diet, with *w¹¹¹⁸* mothers and (bottom row) of crosses of *w¹¹¹⁸* fathers with mutant mothers. The normal *w¹¹¹⁸* IGMR response is also shown (top row). Results are mean \pm SEM of $n = 3$ –8 experiments each with multiple replicates.

(F) ChIP/input signal from modENCODE data sets for leading-edge H3K9me3- and polycomb-dependent genes (red in top panels, blue in bottom panels) in our IGMR offspring embryo RNA-seq. H3K9me3 (top) and H3K27me3 (bottom) enrichment of 12- to 16-hr-old and 16- to 20-hr-old embryos (left) and the difference between the two stages (right). Black lines present the average for all genes.

(G) Box plots of distance to nearest class I and class II insulators. Shown are distances for leading-edge *Su(var)* and *PcG* upregulated genes in our IGMR offspring embryo RNA-seq. Grey boxes represent a control set of equally expressed genes. The boxes indicate the first and third quartiles, and the central line indicates the median. Whiskers extend to the most extreme data point, which is no more than 1.5 times the quartile range.

(H) Distance to nearest class I (orange) and class II (green) insulator plotted according to ranked expression change from IGMR RNA-seq results (high versus medium sugar). Values are sliding window averages of 500 genes.

(I) (Left) H3K9me3 staining of fat body cell nuclei from offspring of medium (30 g/L) and high-sugar (300 g/L) fed fathers. Results are mean \pm SEM of $n = 7$ experiments, each with multiple replicates. (Right) Quantification of fat body cell nucleus H3K9me3 staining.

See also Figure S3.

and 4) and reductions in H3K9me3 immunofluorescence in adult IGMR offspring fat bodies (Figure 6I and Figure S3B) indicate that this chromatin state reprogramming is stable lifelong in the offspring.

A Conserved Signature for Chromatin-State-Associated Phenotypic Variation

More fundamentally, the above data identify a mechanism that directionally controls phenotypic variation within a population.

To probe potential conservation of such processes, we searched for similar signatures in data sets from mouse and man. We examined two murine and three human microarray data sets focusing on adipose tissue from lean and obese individuals, first defining mouse and human ortholog pathways to all Flybase-annotated *Drosophila* *Su(var)*'s and then by using GSEA to test for dysregulation (Figures 7A–7D). Intriguingly, we observed clear signatures of *Su(var)* depletion in obese individuals in two of the most highly genetically controlled human adiposity data

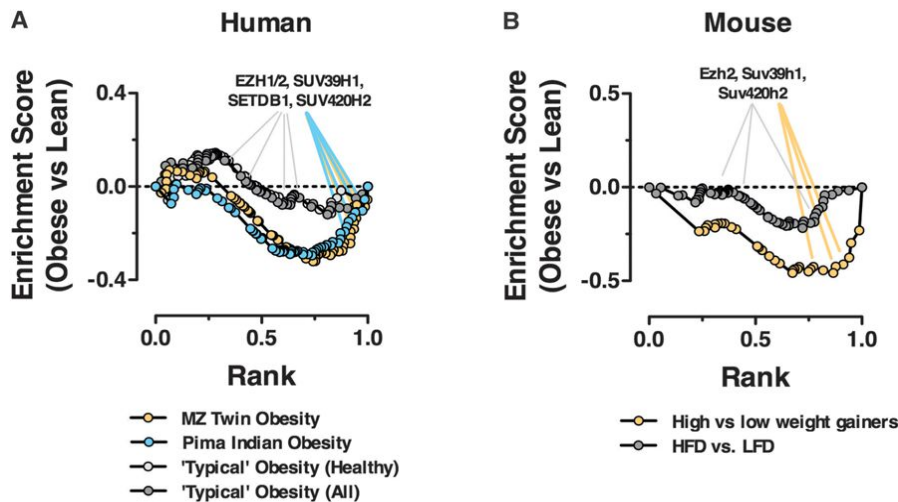


Figure 7. An IGMR Signature Conserved from Fly to Man

GSEA of mouse and human ortholog pathways for all Flybase-annotated *Drosophila* *Su(var)*'s (A) in three human adiposity data sets: 19 obese versus 20 non-obese Pima Indians by Lee et al. (2005); 13 human monozygotic (MZ) twin pairs, each discordant for obesity by Pietiläinen et al. (2008); and a human cohort for "typical" obesity by Klimčáková et al. (2011) and (B) in two murine obesity samples: surgically isolated adipose tissue of future high and low weight gainers from C57BL6/J mice biopsied prior to treatment with high-fat diet by Kozak et al. and diet-induced obesity comparing high- versus low-fat-diet-treated C57BL6/J animals by Voigt et al (2013).

sets available: first in a study of 19 obese versus 20 non-obese Pima Indians (Lee et al., 2005) and then, even more compelling, in a collection of 13 monozygotic twin pairs, each with one normal and one obese co-twin (Pietiläinen et al., 2008) (Figure 7A). Examination of the first figure in the latter study reveals clear evidence also of transcriptome-wide desilencing, with ~5-fold more up- versus downregulated genes in the obese co-twins. Further, similar signatures appear to predict murine obesity susceptibility. In an elegant study, Koza et al. isolated adipose tissue from young C57BL6/J mice prior to treatment with high-fat diet (Koza et al., 2006). Profiling the pretreatment samples from the lowest and highest weight gainers of the 107 animal strong cohort, the authors were able to establish predictive signatures for obesity susceptibility. Reanalyzing these data, we found clear evidence that *Su(var)* pathway depletion predicts obesity susceptibility (Figure 7B).

Leading-edge analysis of all three data sets revealed orthologs of our IGMR defining *Su(var)*3-9, *Setdb1*, *Su(var)*4-20, and *E(z)* regulators as driving the GSEA signal (Figures 7A and 7B, highlighted genes). Importantly, no obvious signatures were observed in an independent "typical" human obesity cohort in which the obesity is most likely driven by assorted genetic factors (Figure 7A) (Klimčáková et al., 2011), nor were they observed in diet-induced obese C57BL6/J mice (Figure 7B) (Voigt et al., 2013). Thus, *Su(var)* suppression characterizes obesity susceptibility on defined human and mouse genetic backgrounds.

These data identify conserved gene signatures for epigenetically defined phenotypic variation from fly to mouse to man.

DISCUSSION

Intergenerational Control of Chromatin State and Obesity

Here, we show that acute dietary interventions, as short as 24 hr, have the capacity to modify F1 offspring phenotype via the male germline. We show that reprogramming occurs in response to dietary manipulations over a physiological range and that phenotypic outcomes require polycomb- and H3K9me3-centric plasticity in spatially and chromatin-state-defined regions of the

genome. The eye color shifts in *w^{m4h}* offspring (Figures 2E, 2G, and 2H) and the reduced fat body H3K9me3 staining in adult IGMR offspring (Figure 6I) supports the conclusions, first, that there are chromatin state changes and, second, that these are stable lifelong. These data are corroborated by selective derepression of *Su(var)*3-9, *SETDB1*, *Su(var)*4-20, and polycomb-sensitive transcripts (Figures 5F–5M); chromatin-state-associated transcriptional rearrangements genome wide (Figures 4F and 6C); selective reprogramming of highly dynamic histone-mark-defined regions (Figure 6F); and the fact that IGMR itself is sensitive to a string of distinct H3K9me3-centric and polycomb mutants (Figures 5A–5E). Although nontrivial, ChIP-seq comparisons of repressive chromatin architecture in mature sperm and multiple defined offspring tissues will be important to establishing the ubiquitousness of these regulatory events and the nature of intergenerational signal itself. These data highlight how acutely sensitive intergenerational control can be to even normal physiological changes, and they identify some of the first genes absolutely required for transmission.

Paternal Diet Regulates Chromatin-Defined Genes in the Germline and Offspring

First categorized simply as heterochromatin versus euchromatin, multiple empirical models now divide the genome into 5 to 51 chromatin states, depending on the analysis (Filion et al., 2010; Kharchenko et al., 2011; Ernst and Kellis, 2010). We find that paternal high sugar increases gene expression preferentially of heterochromatic-embedded genes in embryos. Specifically, these genes are characterized by active deposition of H3K9me3 and H3K27me3, by long distance from class I insulators, and by sensitivity to fully intact expression of *Su(var)*3-9, *Su(var)*4-20, *SetDB1*, *Pc*, and *E(z)*. The data support a model where phenotype has been evolutionarily encoded directly into the chromatin state of relevant loci. Specifically, an abundance of genes important to both cytosolic and mitochondrial metabolism appear to be embedded into H3K9me3- and distinct polycomb-dependent control regions. Indeed, our own GO analysis of the five chromatin colors from Filion et al. (2010) indicate a largely mutually exclusive picture, in which functional pathways

are not randomly distributed across chromatin states (data not shown). Our paternal IGMR data set revealed clear and strong overlaps with pathways of black (lamin-associated) and blue (polycomb) chromatin and included many key metabolic pathways, including *glycolysis*, *TCA cycle*, mitochondrial OxPhos, *chitin*, and *polysaccharide metabolism*, changes that could well prime the system for altered functionality given the appropriate stimulus. Indeed, our paternal IGMR phenotype is a susceptibility to diet-induced obesity and is most readily observable upon high-sugar diet challenge.

Chromatin state coding of functional gene sets would provide a simple mechanism for transgenerational environmental response capable of rewiring even the earliest events of zygotic genome activation. The idea is also consistent with parallel avenues of research already in the literature. rRNA genes, for instance, are not only sensitive to the same *Su(var)*'s but are also known to influence metabolic gene expression and growth (Paredes et al., 2011). Flies with fewer rRNA genes (rDNA) exhibit a phenotype called bobbed (bb), which results in smaller bristles, a reduced growth rate, and a thinner chitinous cuticle (Ritossa et al., 1966). These phenotypes are intriguingly similar to the top GSEA enrichment clusters that we observed for IGMR, namely cell cycle, body morphogenesis, chitin deposition, and metabolism. Interestingly, the very same pathways (chitin synthesis, TCA cycle, carbohydrate, and lipid metabolism) are regulated by nutritional status in third-instar larva (Teleman et al., 2008), suggesting that the paternal IGMR signal acts to prime offspring for metabolic challenge.

Our data support a *trans*-acting mechanism. In the w^{m4h} experiments, male offspring inherited their X chromosome and thus the reporter from their unchallenged mothers, i.e., the reporter allele never encounters the initial signal but is reproducibly reprogrammed. Further, the failure of *Su(var)4-20^{SP}* and *SetDB1¹⁴⁷³* mutants to elicit IGMR responses in their wild-type offspring indicate that wild-type haploid sperm carry the same insufficient reprogramming template as their syncytial mutant counterparts. *cis*- and *trans*-acting mechanisms are not mutually exclusive though. Signals transmitted via paternal chromosomes, though likely transmitted in *cis*, may be manifest via expression of paternal transcripts, which then act in *trans*. Paternal reductions of *Su(var)3-9*, *SetDB1*, and *Hp1*, for instance, would affect the maternal genome in *trans*.

One Genotype, Multiple Paternally Directed Phenotypes

Despite their genetic similarity, isogenic or congeneric animals reared under controlled conditions exhibit measurable variation in essentially all phenotypes. Such variability in genome output is thought to arise largely from probabilistic or chance developmental events in early life (Burga et al., 2011) (review in Whitelaw et al., 2010). Here, we map a mechanism that couples acute paternal feeding and zygotic chromatin state integrity directly to phenotypic output of the next generation. We find that these same signatures predict obesity susceptibility in isogenic mouse and human obesity cohorts. Because acute circadian fluctuations in feeding are essentially constant over evolutionary timescales, they are the perfect mechanistic input upon which a system could evolve to *ensure* defined phenotypic variation within a given population.

EXPERIMENTAL PROCEDURES

Fly Husbandry

Fly stocks were maintained on standard diet at 25°C on a 2 week generation cycle, ensuring a constant ancestral larva density. w^{1118} flies were single-sibling inbred for ten generations and maintained at a fixed fly density for another ten generations before experimental start. Fly strains used: *ln(1) w^{m4h} (w^{m4h})*, *pP(RS5)5-HA-1902 (HA-1902)*, *pP(RS5)5-HA-1992 (HA-1992)*, and *pP(RS5)5-HA-1925 (HA-1925)* and *T(2;3)V21ePlacW (92E)*, *Su(var)3-9⁰⁶*, *Su(var)2-5⁰⁵*, *Su(var)3-1⁰⁴*, *SetDB1¹⁴⁷³*, *Su(var)3-3¹²* from (Phalke et al., 2009). *39C-12* from Sarah Elgin, *Pc³* and *E(z)⁶³* from Leonie Ringrose. PEV lines were single-sibling inbred for ten generations.

Standard diet: Agar 12 g/l, yeast 18 g/l, soy flour 10 g/l, yellow cornmeal 80 g/l, molasses 22 g/l, malt extract 80 g/l, Nipagin 24 g/l, propionic acid 6.25 ml/l. Paternal diet intervention: Agar 12 g/l, yeast 10 g/l, propionic acid 4.5 ml/l, soy flour 30 g/l and white sugar as indicated.

Phenotyping

Body weight of five 7- to 12-day-old males flies was measured on a micro-balance. Wing area determinations were made using ImageJ. Triglycerides (GPO Trinder, Sigma) and glucose and trehalose (Sigma; GAGO-20) were measured on centrifuged cleared lysates from groups of five flies crushed and sonicated in 100 μ l RIPA buffer or TB buffer with or without trehalase (Sigma; T8778-1UN). "Café" assay was performed according to standard procedures. CO₂ production was quantified using a modification of Kucherenko et al. (2011). Eye pigment (A₄₈₀) was measured in centrifuge-cleared sonicates of one or five fly heads in 20/100 μ l RIPA buffer, respectively. Fat body cryosections were fixed for 10 min in 2% formaldehyde in PBS, washed four times for 5 min each in PBS followed by immunofluorescence staining using rabbit anti-H3K9me3 (1:1000, upstate 07-442) and anti-Rabbit Alexa Fluor 488 (1:500; Molecular Probes). Confocal microscopy (LSM 700, Zeiss) analysis used Volocity 5.5 software (Perkin Elmer).

Sperm Dissection

Sperm dissection was modified from Dorus et al. (2006). See additional details in Extended Experimental Procedures.

RNA Sequencing

Trizol-purified RNA was treated with Ribo-Zero (Epicenter) and libraries prepared with a TruSeq stranded kit (Illumina). > 15 million reads per sample were mapped using TopHat v2.0.8, with -G option against the *Drosophila melanogaster* genome (assembly BDGP5, Ensembl release 69). Gene expression values and significantly differentially expressed genes were calculated using Cuffdiff v2.1.1 with upper-quartile normalization and weighting multimapping reads (-N -u options).

Bioinformatic Analysis

Gene set enrichment analysis used GSEA 2.0 or GSEAPreranked with default parameters. Enrichment plots used the Cytoscape plugin Enrichment Map. Analysis of the five chromatin colors used BedTools (2.16.2). For microarray analyses, normalized probe values from the authors were mapped using Ensembl Biomart, and differential analysis against corresponding wild-types were performed using limma in R. Statistically significant was adjusted p value < 0.05 and fold change > 2. Enrichment of chromatin and insulator ChIP-seq data sets from modENCODE used deepTools 1.5.8.1 (Ramírez et al., 2014). Equivalently expressed gene sets were considered as the mean signal of the two genes ranked above and below each gene of interest. Distance to insulators was calculated using BedTools (2.16.2).

Statistical Analysis

Results are presented as means \pm SEM. Statistical tests were performed using one-way ANOVA with a Newman-Keuls posttest. Statistical analysis of chromatin color data sets was a chi-square two-tailed analysis. All statistical analysis was done in with GraphPad Prism, unless otherwise noted.

ACCESSION NUMBERS

The RNA-sequencing data sets reported in this article have been deposited in the NCBI Gene Expression Omnibus and are accessible through GEO series accession number GSE62668.

SUPPLEMENTAL INFORMATION

Supplemental information includes Extended Experimental Procedures, three figures, and four tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2014.11.005>.

AUTHOR CONTRIBUTIONS

A.Ö. and J.A.P. conceived of the study. J.A.P., A.Ö., and A.L. designed the study, supervised all analyses, and wrote the manuscript. A.Ö., A.L., M.W., T.T., and M.D. did all fly work and analyzed the metabolic phenotypes. G.R., M.A., and N.I. contributed to study design and experiments. A.L., C.R., P.M.I., M.D., and A.Ö. performed eating behavior and CO₂ measurements. A.Ö. and A.L. performed immunohistochemistry on fat body cells. A.Ö., A.L., M.W., T.T., U.B., M.S., N.R., and M.R. contributed to generation of the RNaseq data. E.C., S.H., T.V., and L.P. performed the bioinformatic analyses.

ACKNOWLEDGMENTS

This research was supported by the Max-Planck Society, EU (NoE Epigenesis), and the ERC 281641. A.Ö. was supported by Swedish VR K2011-78PK-21893-01-2 and SSMF grants. T.V., L.P., and E.C. were supported by Spanish Ministry grant BFU2011-30246, RYC-2010-07114, Marie Curie European Reintegration Grant “Evo-Chromo,” and the IMPPC. C.R. and P.M.I. were supported by the Champalimaud Foundation, the Human Frontiers Program Project Grant RGP0022/2012, and the Portuguese Foundation for Science and Technology (FCT) grant PTDC/BIA-BCM/118684/2010. P.M.I. is supported by the postdoctoral fellowship SFRH/BPD/79325/2011 from the Foundation for Science and Technology. The authors are grateful to P. Georgiev, S. Raja, T. Manke, J. Longinotto, and T. Lu for important technical and theoretical help; to D. Corona, R. Bodmer, H. Esterbauer, and T. Jenuwein for critical discussions; and to H.J. and J.C.P. for everything.

Received: March 25, 2014

Revised: August 27, 2014

Accepted: October 31, 2014

Published: December 4, 2014

REFERENCES

Anderson, L.M., Riffle, L., Wilson, R., Travlos, G.S., Lubomirski, M.S., and Alvord, W.G. (2006). Preconceptional fasting of fathers alters serum glucose in offspring of mice. *Nutrition* 22, 327–331.

Anway, M.D., Cupp, A.S., Uzumcu, M., and Skinner, M.K. (2005). Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* 308, 1466–1469.

Ashe, A., Sapetschnig, A., Weick, E.M., Mitchell, J., Bagijn, M.P., Cording, A.C., Doebley, A.L., Goldstein, L.D., Lehrbach, N.J., Le Pen, J., et al. (2012). piRNAs can trigger a multigenerational epigenetic memory in the germline of *C. elegans*. *Cell* 150, 88–99.

Blondeau, B., Avril, I., Duchene, B., and Bréant, B. (2002). Endocrine pancreas development is altered in foetuses from rats previously showing intra-uterine growth retardation in response to malnutrition. *Diabetologia* 45, 394–401.

Braunschweig, M., Jagannathan, V., Gutzwiller, A., and Bee, G. (2012). Investigations on transgenerational epigenetic response down the male line in F2 pigs. *PLoS ONE* 7, e30583.

Buescher, J.L., Musselman, L.P., Wilson, C.A., Lang, T., Keleher, M., Baranski, T.J., and Duncan, J.G. (2013). Evidence for transgenerational metabolic programming in *Drosophila*. *Dis. Model. Mech.* 6, 1123–1132.

Burga, A., Casanueva, M.O., and Lehner, B. (2011). Predicting mutation outcome from early stochastic variation in genetic interaction partners. *Nature* 480, 250–253.

Carone, B.R., Fauquier, L., Habib, N., Shea, J.M., Hart, C.E., Li, R., Bock, C., Li, C., Gu, H., Zamore, P.D., et al. (2010). Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. *Cell* 143, 1084–1096.

Daxinger, L., and Whitelaw, E. (2012). Understanding transgenerational epigenetic inheritance via the gametes in mammals. *Nat. Rev. Genet.* 13, 153–162.

Dias, B.G., and Ressler, K.J. (2014). Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nat. Neurosci.* 17, 89–96.

Dorus, S., Busby, S.A., Gerike, U., Shabanowitz, J., Hunt, D.F., and Karr, T.L. (2006). Genomic and functional evolution of the *Drosophila melanogaster* sperm proteome. *Nat. Genet.* 38, 1440–1445.

Dunn, G.A., and Bale, T.L. (2009). Maternal high-fat diet promotes body length increases and insulin insensitivity in second-generation mice. *Endocrinology* 150, 4999–5009.

Ernst, J., and Kellis, M. (2010). Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat. Biotechnol.* 28, 817–825.

Filion, G.J., van Bommel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro, I.J., Kerkhoven, R.M., Bussemaker, H.J., and van Steensel, B. (2010). Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* 143, 212–224.

Fullston, T., Ohlsson Teague, E.M., Palmer, N.O., DeBlasio, M.J., Mitchell, M., Corbett, M., Print, C.G., Owens, J.A., and Lane, M. (2013). Paternal obesity initiates metabolic disturbances in two generations of mice with incomplete penetrance to the F2 generation and alters the transcriptional profile of testis and sperm microRNA content. *FASEB J.* 27, 4226–4243.

Gapp, K., Jawaid, A., Sarkies, P., Bohacek, J., Pelczar, P., Prados, J., Farinelli, L., Miska, E., and Mansuy, I.M. (2014). Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. *Nat. Neurosci.* 17, 667–669.

Gniuli, D., Calcagno, A., Caristo, M.E., Mancuso, A., Macchi, V., Mingrone, G., and Vettor, R. (2008). Effects of high-fat diet exposure during fetal life on type 2 diabetes development in the progeny. *J. Lipid Res.* 49, 1936–1945.

Goodliffe, J.M., Cole, M.D., and Wieschaus, E. (2007). Coordinated regulation of Myc trans-activation targets by Polycomb and the Trithorax group protein Ash1. *BMC Mol. Biol.* 8, 40.

Greer, E.L., Maures, T.J., Ucar, D., Hauswirth, A.G., Mancini, E., Lim, J.P., Benayoun, B.A., Shi, Y., and Brunet, A. (2011). Transgenerational epigenetic inheritance of longevity in *Caenorhabditis elegans*. *Nature* 479, 365–371.

Jimenez-Chillaron, J.C., Isganaitis, E., Charalambous, M., Gesta, S., Pentinat-Pelegrin, T., Faucette, R.R., Otis, J.P., Chow, A., Diaz, R., Ferguson-Smith, A., and Patti, M.E. (2009). Intergenerational transmission of glucose intolerance and obesity by in utero undernutrition in mice. *Diabetes* 58, 460–468.

Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T., et al. (2011). Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* 471, 480–485.

Kiani, J., Grandjean, V., Liebers, R., Tuorto, F., Ghanbarian, H., Lyko, F., Cuzin, F., and Rassoulzadegan, M. (2013). RNA-mediated epigenetic heredity requires the cytosine methyltransferase Dnmt2. *PLoS Genet.* 9, e1003498.

Klimáková, E., Roussel, B., Márquez-Quiriones, A., Kováčová, Z., Kováčiková, M., Combes, M., Siklová-Vitková, M., Hejnová, J., Srámková, P., Bouloumié, A., et al. (2011). Worsening of obesity and metabolic status yields similar molecular adaptations in human subcutaneous and visceral adipose tissue: decreased metabolism and increased immune response. *J. Clin. Endocrinol. Metab.* 96, E73–E82.

Koza, R.A., Nikonova, L., Hogan, J., Rim, J.S., Mendoza, T., Faulk, C., Skaf, J., and Kozak, L.P. (2006). Changes in gene expression foreshadow diet-induced obesity in genetically identical mice. *PLoS Genet.* 2, e81.

- Kucherenko, M.M., Marrone, A.K., Rishko, V.M., Magliarelli, Hde.F., and Shcherbata, H.R. (2011). Stress and muscular dystrophy: a genetic screen for dystroglycan and dystrophin interactors in *Drosophila* identifies cellular stress response components. *Dev. Biol.* 352, 228–242.
- Lee, Y.H., Nair, S., Rousseau, E., Allison, D.B., Page, G.P., Tataranni, P.A., Bogardus, C., and Permana, P.A. (2005). Microarray profiling of isolated abdominal subcutaneous adipocytes from obese vs non-obese Pima Indians: increased expression of inflammation-related genes. *Diabetologia* 48, 1776–1783.
- Lee, H.C., Gu, W., Shirayama, M., Youngman, E., Conte, D., Jr., and Mello, C.C. (2012). *C. elegans* piRNAs mediate the genome-wide surveillance of the germline transcripts. *Cell* 150, 78–87.
- Lundberg, L.E., Stenberg, P., and Larsson, J. (2013). HP1a, Su(var)3-9, SETDB1 and POF stimulate or repress gene expression depending on genomic position, gene length and expression pattern in *Drosophila melanogaster*. *Nucleic Acids Res.* 41, 4481–4494.
- Merico, D., Isserlin, R., Stueker, O., Emili, A., and Bader, G.D. (2010). Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS ONE* 5, e13984.
- Morgan, H.D., Sutherland, H.G., Martin, D.I., and Whitelaw, E. (1999). Epigenetic inheritance at the agouti locus in the mouse. *Nat. Genet.* 23, 314–318.
- Nègre, N., Brown, C.D., Shah, P.K., Kheradpour, P., Morrison, C.A., Henikoff, J.G., Feng, X., Ahmad, K., Russell, S., White, R.A., et al. (2010). A comprehensive map of insulator elements for the *Drosophila* genome. *PLoS Genet.* 6, e1000814.
- Nègre, N., Brown, C.D., Ma, L., Bristow, C.A., Miller, S.W., Wagner, U., Kheradpour, P., Eaton, M.L., Loriaux, P., Sealfon, R., et al. (2011). A cis-regulatory map of the *Drosophila* genome. *Nature* 471, 527–531.
- Ng, S.F., Lin, R.C., Laybutt, D.R., Barres, R., Owens, J.A., and Morris, M.J. (2010). Chronic high-fat diet in fathers programs β -cell dysfunction in female rat offspring. *Nature* 467, 963–966.
- Ozanne, S.E., Wang, C.L., Dorling, M.W., and Petry, C.J. (1999). Dissection of the metabolic actions of insulin in adipocytes from early growth-retarded male rats. *J. Endocrinol.* 162, 313–319.
- Paredes, S., Branco, A.T., Hartl, D.L., Maggert, K.A., and Lemos, B. (2011). Ribosomal DNA deletions modulate genome-wide gene expression: “rDNA-sensitive” genes and natural variation. *PLoS Genet.* 7, e1001376.
- Patti, M.E. (2013). Intergenerational programming of metabolic disease: evidence from human populations and experimental animal models. *Cell. Mol. Life Sci.* 70, 1597–1608.
- Phalke, S., Nickel, O., Walluscheck, D., Hortig, F., Onorati, M.C., and Reuter, G. (2009). Retrotransposon silencing and telomere integrity in somatic cells of *Drosophila* depends on the cytosine-5 methyltransferase DNMT2. *Nat. Genet.* 41, 696–702.
- Pietiläinen, K.H., Naukkarinen, J., Rissanen, A., Saharinen, J., Ellonen, P., Keränen, H., Suomalainen, A., Götz, A., Suortti, T., Yki-Järvinen, H., et al. (2008). Global transcript profiles of fat in monozygotic twins discordant for BMI: pathways behind acquired obesity. *PLoS Med.* 5, e51.
- Radford, E.J., Ito, M., Shi, H., Corish, J.A., Yamazawa, K., Isganaitis, E., Seisenberger, S., Hore, T.A., Reik, W., et al. (2014). In utero undernourishment perturbs the adult sperm methylome and intergenerational metabolism. *Science* 345. Published online July 10, 2014. <http://dx.doi.org/10.1126/science.1255903>.
- Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* 42 (Web Server issue), W187–W191.
- Rando, O.J. (2012). Daddy issues: paternal effects on phenotype. *Cell* 151, 702–708.
- Rassoulzadegan, M., Grandjean, V., Gounon, P., Vincent, S., Gillot, I., and Cuzin, F. (2006). RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature* 441, 469–474.
- Rechavi, O., Houriz-Ze’evi, L., Anava, S., Goh, W.S., Kerk, S.Y., Hannon, G.J., and Hobert, O. (2014). Starvation-induced transgenerational inheritance of small RNAs in *C. elegans*. *Cell* 158, 277–287.
- Ritossa, F.M., Atwood, K.C., and Spiegelman, S. (1966). A molecular explanation of the bobbed mutants of *Drosophila* as partial deficiencies of “ribosomal” DNA. *Genetics* 54, 819–834.
- Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., and Jenuwein, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.* 18, 1251–1262.
- Seong, K.H., Li, D., Shimizu, H., Nakamura, R., and Ishii, S. (2011). Inheritance of stress-induced, ATF-2-dependent epigenetic change. *Cell* 145, 1049–1061.
- Shirayama, M., Seth, M., Lee, H.C., Gu, W., Ishidate, T., Conte, D., Jr., and Mello, C.C. (2012). piRNAs initiate an epigenetic memory of nonself RNA in the *C. elegans* germline. *Cell* 150, 65–77.
- Skorupa, D.A., Dervisevendic, A., Zwiener, J., and Pletcher, S.D. (2008). Dietary composition specifies consumption, obesity, and lifespan in *Drosophila melanogaster*. *Aging Cell* 7, 478–490.
- Teleman, A.A., Hietakangas, V., Sayadian, A.C., and Cohen, S.M. (2008). Nutritional control of protein biosynthetic capacity by insulin via Myc in *Drosophila*. *Cell Metab.* 7, 21–32.
- Voigt, A., Agnew, K., van Schothorst, E.M., Keijer, J., and Klaus, S. (2013). Short-term, high fat feeding-induced changes in white adipose tissue gene expression are highly predictive for long-term changes. *Mol. Nutr. Food Res.* 57, 1423–1434.
- Waterland, R.A., Dolinoy, D.C., Lin, J.R., Smith, C.A., Shi, X., and Tahiliani, K.G. (2006). Maternal methyl supplements increase offspring DNA methylation at Axin Fused. *Genesis* 44, 401–406.
- Whitelaw, N.C., Chong, S., and Whitelaw, E. (2010). Tuning in to noise: epigenetics and intangible variation. *Dev. Cell* 19, 649–650.
- Wolff, G.L., Kodell, R.L., Moore, S.R., and Cooney, C.A. (1998). Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a mice. *FASEB J.* 12, 949–957.
- Xing, Y., Shi, S., Le, L., Lee, C.A., Silver-Morse, L., and Li, W.X. (2007). Evidence for transgenerational transmission of epigenetic tumor susceptibility in *Drosophila*. *PLoS Genet.* 3, 1598–1606.

Uridylation by TUT4 and TUT7 Marks mRNA for Degradation

Jaechul Lim,^{1,2,4} Minju Ha,^{1,2,4} Hyeshik Chang,^{1,2,4} S. Chul Kwon,^{1,2} Dharendra K. Simanshu,³ Dinshaw J. Patel,³ and V. Narry Kim^{1,2,*}

¹Center for RNA Research, Institute for Basic Science, Seoul 151-742, Korea

²School of Biological Sciences, Seoul National University, Seoul 151-742, Korea

³Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA

⁴Co-first author

*Correspondence: narrykim@snu.ac.kr

<http://dx.doi.org/10.1016/j.cell.2014.10.055>

SUMMARY

Uridylation occurs pervasively on mRNAs, yet its mechanism and significance remain unknown. By applying TAIL-seq, we identify TUT4 and TUT7 (TUT4/7), also known as ZCCHC11 and ZCCHC6, respectively, as mRNA uridylation enzymes. Uridylation readily occurs on deadenylated mRNAs in cells. Consistently, purified TUT4/7 selectively recognize and uridylate RNAs with short A-tails (less than ~25 nt) *in vitro*. PABPC1 antagonizes uridylation of polyadenylated mRNAs, contributing to the specificity for short A-tails. In cells depleted of TUT4/7, the vast majority of mRNAs lose the oligo-U-tails, and their half-lives are extended. Suppression of mRNA decay factors leads to the accumulation of oligo-uridylated mRNAs. In line with this, microRNA induces uridylation of its targets, and TUT4/7 are required for enhanced decay of microRNA targets. Our study explains the mechanism underlying selective uridylation of deadenylated mRNAs and demonstrates a fundamental role of oligo-U-tail as a molecular mark for global mRNA decay.

INTRODUCTION

RNA tailing (nontemplated nucleotide addition to the 3' end of RNA) is one of the most frequent types of RNA modification, with a deep evolutionary root and diverse molecular functions. In bacteria, adenylation of mRNA triggers RNA degradation whereas polyadenylation in eukaryotes increases the stability and translatability of mRNA (Dreyfus and Régnier, 2002). Tailing is catalyzed by a group of template-independent ribonucleotidyl transferases that contain DNA polymerase β -like nucleotidyl transferase domain (Aravind and Koonin, 1999). Apart from canonical poly(A) polymerases (PAPs) that generate poly(A) tail of mRNA, many noncanonical PAPs have been described from fission yeast to human (Martin and Keller, 2007; Norbury, 2013). Because some noncanonical PAPs catalyze uridylation instead of adenylation, noncanonical PAPs are also called terminal uridylyl transferases (TUTases or TUTs). Some PAPs/TUTs

have more relaxed nucleotide specificity and carry out both uridylation and adenylation. Humans have seven noncanonical PAPs/TUTs with distinct substrate specificity and subcellular localization.

Uridylation of mRNA was initially noticed at the 3' ends of miRNA-directed cleavage products in *Arabidopsis* and mammalian cells (Shen and Goodman, 2004). U-tails were also detected on human replication-dependent histone mRNAs that lack a poly(A) tail (Mullen and Marzluff, 2008). Histone mRNAs are uridylated and degraded at the end of S phase or upon inhibition of DNA replication (Mullen and Marzluff, 2008). TUT4 (ZCCHC11) was reported to catalyze histone mRNA uridylation (Schmidt et al., 2011; Su et al., 2013), although two other TUTs (TUT1/MTPAP/PAPD1 and TUT3/PAPD5/TRF4-2) were proposed in an earlier study (Mullen and Marzluff, 2008). Uridylation induces rapid decay of histone mRNA through both the 5'–3' degradation by XRN1, DCP2, and LSM1 and the 3'–5' degradation by exosome and ERI1 (3'hExo) (Hoefig et al., 2013; Mullen and Marzluff, 2008; Slevin et al., 2014).

Interestingly, uridylation occurs not only on poly(A)-lacking mRNAs but also on poly(A)⁺ mRNAs, as shown first with the actin (*act1*) mRNA in fission yeast *Schizosaccharomyces pombe* (Rissland et al., 2007). When six mRNAs were examined by circularized rapid amplification of cDNA ends (cRACE) technique, all of them were found to bear short U-tails (usually one or two uridines) at the end of poly(A) tails albeit at varying frequencies, indicating that mRNA uridylation may be widespread in fission yeast (Rissland and Norbury, 2009). The stability of the *urg1* mRNA increased in a mutant lacking Cid1 which is one of the TUTs in fission yeast (Rissland et al., 2007; Rissland and Norbury, 2009). The uridylation frequency was enhanced in mutants defective of deadenylase and decapping enzyme (*ccr4Δ* and *dcp1-ts*). Based on these results, it was proposed that uridylation and deadenylation may act redundantly to induce decapping. A more recent study showed that *Arabidopsis* mRNAs are also subject to uridylation (Sement et al., 2013). Short uridyl residues (1–2 uridines) were detected on deadenylated, decapped mRNAs. The Cid1 homolog URT1 is required for uridylation. But, curiously, *URT1* mutation did not have a major impact on mRNA turnover and instead inhibited trimming of mRNA from the 3' end (Sement et al., 2013), implying that uridylation may be necessary to establish the directionality (5'–3') rather than to control the rate of mRNA decay. Therefore, although these

observations are intriguing, it was unclear if uridylation has a conserved function across species and whether animal poly(A)⁺ mRNAs are also uridylated. In addition, because previous studies examined a few individual mRNAs by RACE and small-scale cloning, it remained to be tested whether or not uridylation occurs globally and if the observed changes in uridylation and poly(A) length are statistically significant.

To investigate tail structures at the genomic scale, we recently developed a method called TAIL-seq that deep-sequences the 3' most fragments of RNAs (Chang et al., 2014b). The TAIL-seq protocol begins with removal of abundant noncoding RNAs such as rRNA, tRNA, small nuclear RNA (snRNA), and small nucleolar RNA (snoRNA) by affinity-based depletion and size fractionation. To avoid any bias against unconventional tails, TAIL-seq does not use splint ligation or oligo-d(T) enrichment. The resulting RNA sample enriched with mRNA is subsequently ligated to the 3' adaptor that contains biotin residues. Following partial fragmentation, the 3' most fragments are purified using streptavidin beads and ligated to the 5' adaptor. Paired-end sequencing of the cDNA library yields 51 nt from the 5' terminus of the fragment (to identify the transcript) and 231 nt from the 3' terminus (to examine the tail sequences).

TAIL-seq provided us with a unique opportunity to investigate poly(A) tail length and additional 3' modifications simultaneously at the genomic scale. Surprisingly, we found that the vast majority of mRNAs are subject to uridylation in mammals. Over 85% of mRNAs are terminally uridylated at a frequency of higher than 1% in both NIH 3T3 and HeLa cells (Chang et al., 2014b). Interestingly, U-tails are found mainly on mRNAs with short A-tails (less than ~25 nt), indicating that uridylation may occur following deadenylation. We further detected a negative correlation between uridylation frequency and mRNA half-life, suggesting a role of uridylation in general mRNA decay.

Current model for eukaryotic mRNA decay pathway is mainly based on the pioneering genetic and biochemical studies in *Saccharomyces cerevisiae* (Garneau et al., 2007; Houseley and Tollervey, 2009; Norbury, 2013; Parker and Song, 2004). Decay is generally initiated by deadenylation that is mediated by multiple deadenylases such as the Pan2-Pan3 complex and the Ccr4-Not complex. Subsequently, deadenylated mRNAs are subject to either of two major decay pathways. In the 5'–3' decay pathway, the Lsm1–7 complex binds to the 3' end of deadenylated mRNA and recruits the decapping complex (Dcp1/2) that removes 5' cap structure. Subsequently, 5' monophosphate-dependent exoribonuclease, Xrn1, digests mRNA processively. From the opposite orientation, a multisubunit exosome complex degrades deadenylated mRNAs from the 3' end. This model seems to apply generally to most, if not all, eukaryotic species. However, *S. cerevisiae* is unusual among eukaryotes in that it does not have any known TUT homolog with uridylation activity and that mRNAs in *S. cerevisiae* do not carry terminal U-tails (Norbury, 2013). Thus, the current model for mRNA decay, particularly in mammals, may need to be revised to incorporate the recent findings of pervasive uridylation (Lee et al., 2014).

In this study, we aimed to identify enzyme(s) that catalyze mRNA uridylation in mammals and understand the significance of uridylation in the mRNA decay pathway. We discover TUT4 and TUT7 as uridylyl transferases for poly(A)⁺ mRNAs in humans

and delineate in detail the action mechanism and molecular function of uridylation in the mRNA decay pathway. Based on these results, we propose a revised model for general mRNA decay in mammals.

RESULT

TUT4 and TUT7 Catalyze mRNA Uridylation

In order to identify enzyme(s) responsible for mRNA uridylation, we took a candidate approach by depleting seven human TUTases (Figure S1A available online). Because TUT2 (also known as GLD2 and PAPD4), TUT4 (ZCCHC11), and TUT7 (ZCCHC6) act redundantly in mono-uridylation of precursor of let-7 (pre-let-7) (Heo et al., 2012), we knocked down TUTases in two subgroups (TUT1/3/5/6 and TUT2/4/7) by transfecting siRNA mixtures into HeLa cells (Figure S1B) and carried out TAIL-seq (Figure 1A). Overall frequency of uridylation was quantified by dividing the read number of terminally uridylated mRNAs by that of total mRNAs. Because short A-tails are preferentially uridylated (Chang et al., 2014b), uridylation frequency in short A-tail range (5–25 nt) is shown in Figure 1. Interestingly, when TUT2/4/7 were depleted, terminal uridylation was significantly reduced while RNAi of TUT1/3/5/6 did not affect uridylation.

To narrow down on individual TUTases, we generated knockout HeLa cell lines using TALENs (transcription activator-like effector nucleases) against the genes coding TUT2, TUT4, or TUT7 proteins (Figure S1C). We observed a modest decrease of uridylation in both *TUT4* and *TUT7* knockout cells, but not in *TUT2* knockout cells (Figure 1B). Repeated attempts to generate double knockout of *TUT4* and *TUT7* by utilizing the TALEN and CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9) systems have failed (Figure S1D). Although genomic deletion was effectively introduced by the nucleases, mutant clones disappeared during clonal selection processes (Figure S1D), which indicates that the combined activity of TUT4 and TUT7 is essential for cell viability.

Of note, previous studies have shown that TUT4 and TUT7 are highly similar in their domain organization and activity in pre-miRNA uridylation (Heo et al., 2012; Liu et al., 2014; Thornton et al., 2012). Thus, TUT4 and TUT7 (TUT4/7) may act redundantly in mRNA uridylation as well as in pre-miRNA uridylation. To confirm this notion, we carried out simultaneous transient RNAi against TUT4/7 by transfecting siRNAs (Figures 1C and S1E). The TUT4/7 knockdown cells looked largely normal and proliferated at a modestly reduced rate with a slight increase of apoptosis after 4 days of siRNA treatment (Figures S1F and S1G). Under this condition, uridylation of mRNA was significantly reduced when both TUT4 and TUT7 are depleted (Figure 1C). Oligo-uridylation (≥ 2 U) was more sensitive to TUT4/7 knockdown than mono-uridylation was (3.71-fold and 1.36-fold decrease, respectively), suggesting that a relatively high level of TUT4/7 may be required to generate oligo-U-tails on mRNA.

Gene-level analyses revealed that the majority of mRNA species (638 out of 746 genes, 85.5%) are decreased in uridylation following TUT4/7 knockdown ($p = 7.69 \times 10^{-100}$, one-tailed Mann-Whitney U test) (Figure 1D; Table S1). This result strongly indicates that TUT4/7 uridylate most, if not all, mRNAs. Figure 1E presents 21 most abundant mRNAs as examples, the majority of

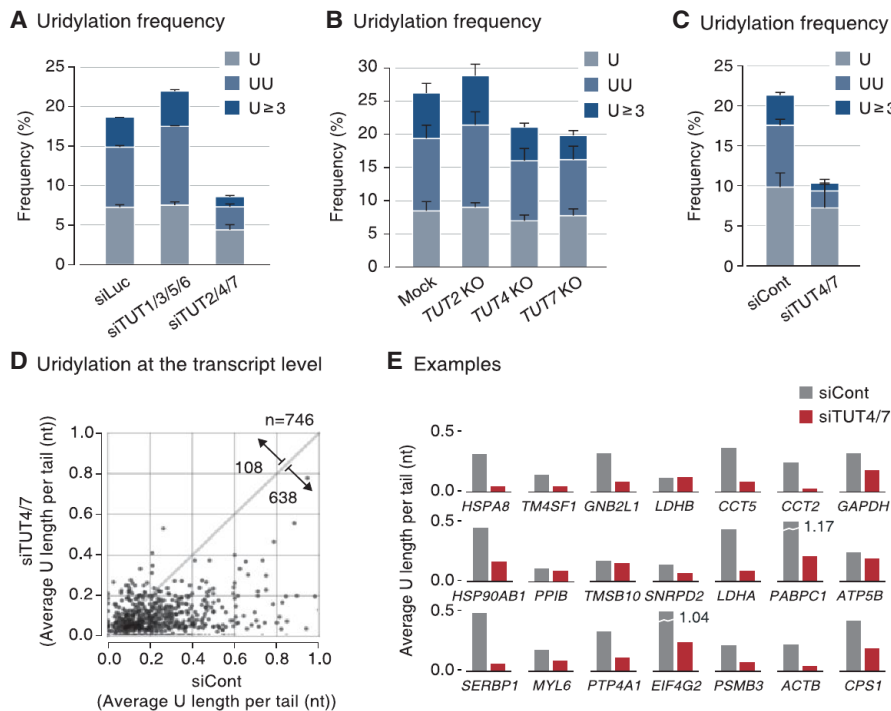


Figure 1. TUT4 and TUT7 Are Required for mRNA Uridylation in Human Cells

(A) Uridylation frequency measured by small-scale TAIL-seq (with Illumina MiSeq) following RNAi of the indicated genes. Frequency (y axis) is the fraction of uridylated reads among the total number of mRNA reads with short poly(A) tail (5–25 nt). Light blue refers to mono-uridylation (U), blue indicates di-uridylation (UU), and dark blue represents ≥ 3 uridines ($U \geq 3$). Uridylation frequency significantly decreased in siTUT2/4/7 ($p = 0.0378$ for U; 0.0388 for UU; 0.0201 for $U \geq 3$ by one-tailed t test). Error bar represents SEM from two biologically independent replicates ($n = 2$).

(B) Uridylation frequency of mRNAs with short poly(A) tails (5–25 nt) measured by small-scale TAIL-seq in knockout HeLa cell lines. Uridylation frequency was reduced modestly in *TUT4* and *TUT7* knockout cells ($p = 0.109$ for U, 0.0273 for UU, 0.142 for $U \geq 3$ of *TUT4* KO; $p = 0.150$ for U, 0.00685 for UU, 0.0713 for $U \geq 3$ of *TUT7* KO by one-tailed t test). Error bar represents SEM from two replicates ($n = 2$).

(C) Uridylation frequency of mRNAs with short poly(A) tails (5–25 nt) measured by TAIL-seq following simultaneous *TUT4* and *TUT7* knockdown (siTUT4/7). Uridylation was reduced when both *TUT4* and *TUT7* were depleted ($p = 0.0941$ for U, 0.00922 for UU, 0.0105 for

$U \geq 3$; one-tailed t test). Error bar represents SEM from three biological replicates ($n = 3$).

(D) Changes in uridylation of individual mRNA species upon *TUT4/7* knockdown. “Average U length per tail” (y axis) is the sum of the number of all uridines on short A-tails (5–25 nt) divided by the total number of reads with short A-tails. Note that unlike “uridylation frequency,” average U length per tail weighs every uridine in oligo-U-tails. Each dot represents a transcript with ≥ 15 reads in both samples. Uridylation was significantly decreased following *TUT4/7* knockdown ($p = 7.69 \times 10^{-100}$, one-tailed Mann-Whitney U test). The full list is shown in Table S1.

(E) Examples of gene-level uridylation changes. Twenty-one most abundant mRNAs (not including ribosomal protein mRNAs and histone mRNAs) are shown in the order of mRNA abundance.

See also Figure S1.

which are reduced in uridylation upon *TUT4/7* knockdown. Two biological replicate experiments showed a comparable decrease of uridylation (Figure S1H).

Histone mRNAs that lack poly(A) tails are also uridylated and their uridylation is dependent modestly on *TUT4/7*, but not on *TUT1/2/3/5/6* (data not shown). However, poly(A)[−] histone mRNAs were excluded from our current data analyses because we used nonsynchronous cell population for our experiments, and it is known that uridylation of histone mRNA occurs specifically at the end of S phase (Mullen and Marzluff, 2008; Schmidt et al., 2011; Su et al., 2013). It would be more appropriate to investigate histone mRNAs using synchronous cells in future studies.

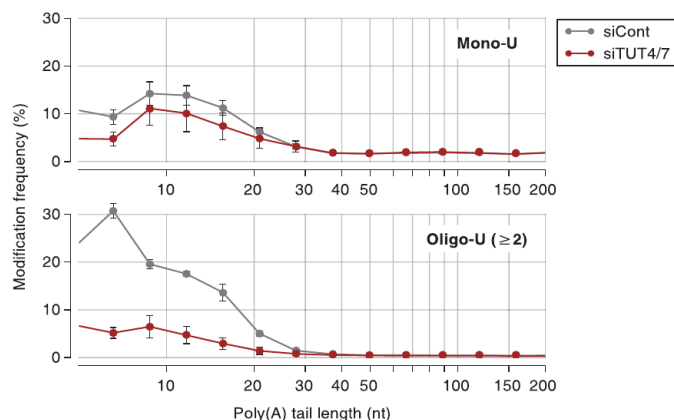
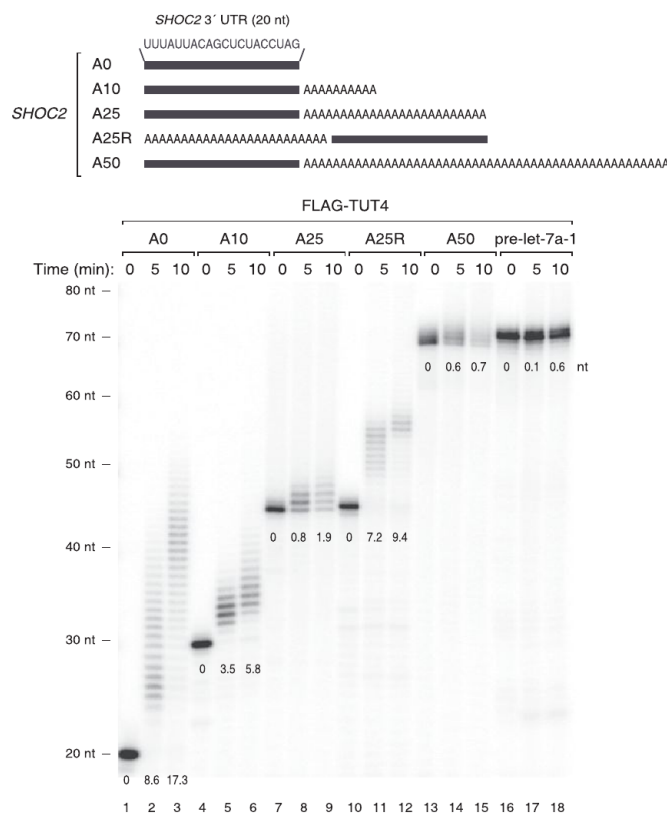
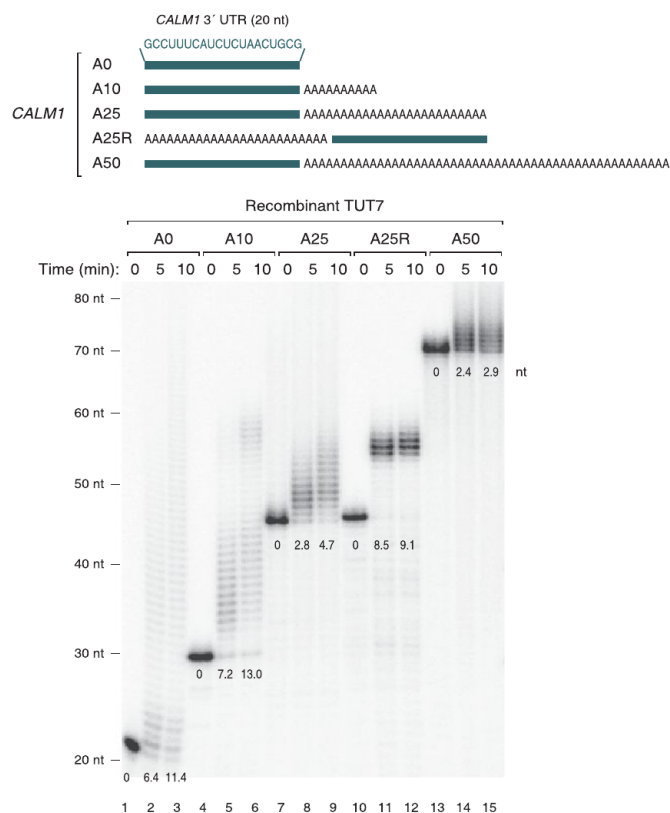
TUT4/7 Selectively Oligo-Uridylate mRNAs with Short A-Tails In Vivo and In Vitro

It is intriguing that uridylation occurs preferentially on shortened A-tails in plants and animals (Chang et al., 2014b; Sement et al., 2013). Figure 2A shows the distribution of U-tails over different lengths of A-tails in HeLa cells. The frequency of uridylation on the transcripts with a short A-tails (5–25 nt) is higher than that on the rest (A-tails of >25 nt), especially when only oligo-U (≥ 2 U) is counted. Note that mRNAs with A-tails of shorter than 5 nt were excluded from this analysis as it is sometimes diffi-

cult to distinguish them from genomic A-rich sequences in 3' UTR. When *TUT4/7* were depleted, uridylation on short A-tails was selectively reduced (especially for oligo-U), indicating that *TUT4/7* are responsible for the specific uridylation of short A-tails (Figure 2A).

To understand the mechanism underlying such strong association with A-tail length, we performed in vitro uridylation assays using immunopurified full-length TUTases (Figures 2B and S2A–S2C). Substrate RNAs were chemically synthesized to contain heterogenous sequences (the last 20 nt from the *SHOC2* 3' UTR) linked to A-tails of various lengths (0, 10, 25, and 50 nt) at the 3' end (Figure 2B). We also used a “swapped” control (A25R) that has a 25 nt A segment at the 5' side of the *SHOC2* 3' UTR such that the RNA is identical to *SHOC2*-A25 (A25) in the overall length and base composition, but lacks an A-tail at the 3' end (Figure 2B).

Interestingly, RNAs with no tail (A0) or a short A-tail (A10) were oligo-uridylated efficiently by *TUT4* under the condition where pre-let-7a-1 is mono-uridylated weakly (Figure 2B). A25 and A50 were less efficiently uridylated than A0 and A10 were. The A25R RNA was a much better substrate than the A25 was, indicating that it is the 3' A-tail length (not the overall RNA length) that is measured by *TUT4* (Figure 2B). Comparable results were obtained with full-length *TUT7* protein (Figures S2A and S2B), again

A Uridylation frequency**B** In vitro uridylation assay with immunopurified TUT4**C** In vitro uridylation assay with recombinant TUT7**Figure 2. Short A-Tails Are Selectively Uridylated by TUT4 and TUT7**

(A) Distribution of mono-uridylation (top) and oligo-uridylation (bottom) according to the length of poly(A) tails. Poly(A) tail lengths from 5 nt to 231 nt are pooled into equal-width bins in the logarithmic scale (base 2) (x axis). The left edges (inclusive) of bins are 5, 7, 9, 12, 15, 21, 28, 38, 50, 67, 89, 119, 159, and 212 nt. Uridylation frequency (y axis) indicates the percentage of uridylated reads within each poly(A) tail size range. Error bar represents SEM (n = 3).

(B) Top: illustration of chemically synthesized RNA substrates. Grey bars represent the last 20 nt of *SHOC2* 3' UTR and "A" indicates an adenosine. Bottom: in vitro uridylation assay using immunopurified FLAG-TUT4. RNA (0.45 nM) was used in each reaction. The products were resolved on 6% polyacrylamide sequencing gel containing 7 M urea. The average length of uridylation is shown below each band. See Extended Experimental Procedures for quantification method.

(C) Top: illustration of chemically synthesized RNA substrates. Green bars represent the last 20 nt of *CALM1* 3' UTR and "A" indicates an adenosine. Bottom: in vitro uridylation assay using recombinant TUT7 C-terminal fragment (951–1,495 aa) purified from *E. coli*. RNA (0.45 nM) and 14 nM of recombinant TUT7 were used in each reaction. Extension products were resolved on 6% polyacrylamide sequencing gel containing 7 M urea. The average length of uridylation was quantified as in (B). See also Figure S2.

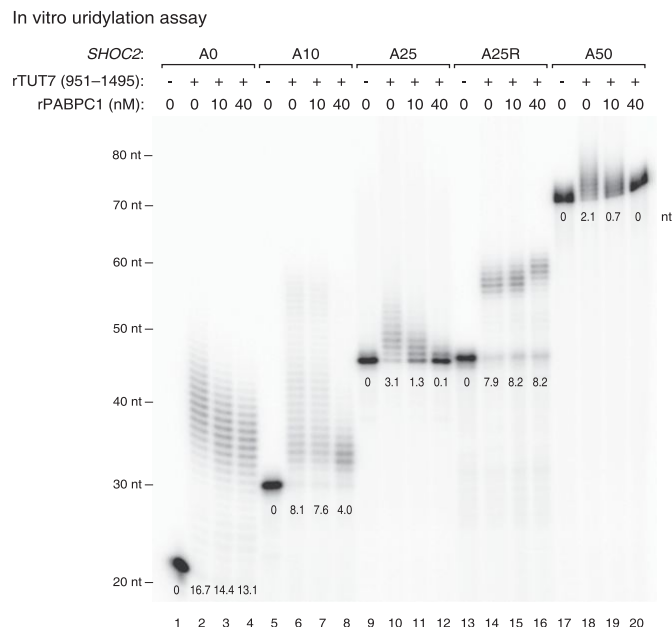


Figure 3. PABP Inhibits Uridylation of Polyadenylated mRNA

In vitro uridylation assay by using recombinant TUT7 (951–1,495 aa) with a varying concentration of recombinant PABPC1 (0, 10, or 40 nM). 0.45 nM of RNA and 160 nM of recombinant TUT7 (rTUT7) were used in the reaction. Extension products were resolved on 6% polyacrylamide sequencing gel containing 7 M urea. The average length of uridylation was quantified as described in Extended Experimental Procedures and shown below each band. See also Figure S3.

demonstrating that these two related enzymes are functional paralogs. The U-tail length in Figures S2A and S2B was overall shorter than those in Figure 2B because the amount of immunoprecipitated TUT7 was smaller than that of TUT4 in Figure 2B (data not shown).

We also prepared recombinant TUT7 protein (951–1,495 aa) from *Escherichia coli* and used the fragment for in vitro uridylation assay (Figure S2C). Apart from the *SHOC2* RNAs (Figure S2D), we synthesized and tested another series of RNAs based on the *CALM1* 3' UTR sequences (Figure 2C). The purified protein fragment was fully capable of carrying out uridylation in an A-tail length-dependent manner with both RNAs (Figures 2C and S2D, see below). Thus, the C-terminal half of TUT7 is sufficient to recognize and uridylate single-stranded RNAs with a short A-tails (less than ~25 nt), in a 3' UTR sequence-independent manner. These results suggest that TUT4/7 possess an intrinsic ability to measure the 3' terminal A length and avoid uridylation of long A-tails.

PABP Suppresses Uridylation of Poly(A)⁺ mRNA

As poly(A)⁺ mRNAs are associated with poly(A) binding protein (PABP) in cells, we asked if PABP has an influence on mRNA uridylation. It was previously shown that PABP preferentially interacts with poly(A) or A-rich sequences (Eliseeva et al., 2013). The binding affinity increases as the A stretch gets longer (Eliseeva et al., 2013; Khanam et al., 2006; Kühn and Pieler, 1996; Sachs et al., 1987). Full-length PABP occupies an ~25 nt A-tail

as determined by nuclease digestion assay (Baer and Kornberg, 1983; Eliseeva et al., 2013). In order to test an effect of PABP on uridylation, we carried out in vitro uridylation assays in the presence of recombinant PABPC1 (Figure 3). When PABPC1 was added to RNA, uridylation of RNAs with long poly(A) tail (A25 and A50) was suppressed even at a low concentration of PABPC1 (10 nM) while those with no or short A-tail (A0, A10, and A25R) remained largely unaffected (Figure 3). This result suggests that PABPC1 binds preferentially to long poly(A) tails and protects them from TUT4/7 and thereby enhances the selectivity of uridylation according to poly(A) tail length.

Taken together, our results suggest that the strict dependence on the A-tail length observed in vivo may be determined by the combination of two factors: (1) the intrinsic ability of TUT4/7 to measure poly(A) stretch (Figure 2), and (2) the protective activity of PABP (Figure 3).

As deadenylation is thought to occur mainly in the cytoplasm, we examined the localization of TUT4/7 by western blotting. The TUT4 and TUT7 proteins are mainly localized in the cytoplasm (Figure S3). Thus, TUT4/7 may function mainly in the metabolism of cytoplasmic, deadenylated mRNAs.

Uridylation Facilitates Global mRNA Decay

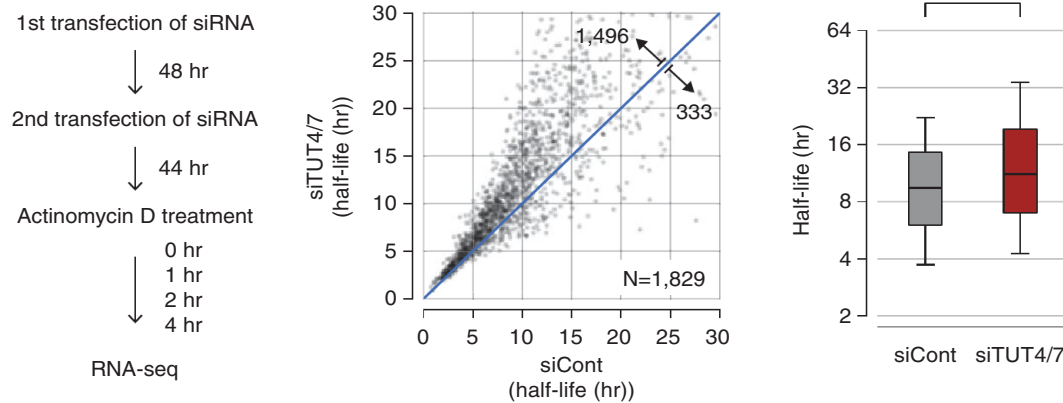
To understand the functional consequences of uridylation, we measured mRNA half-life in HeLa cells with or without TUT4/7 knockdown (Figure 4A). mRNA levels were determined by RNA-seq at 0, 1, 2, and 4 hr after actinomycin D treatment that blocks transcription. To avoid any bias from tail length variation, we omitted the oligo-dT enrichment step and instead used RiboZero to remove abundant rRNAs prior to cDNA library construction. We could measure turnover rates of 1,829 mRNAs. In TUT4/7-depleted cells, the majority of mRNAs (1,426 out of 1,829 [78.0%]) showed increase stability (Figure 4A, left panel; Table S2). Half-lives were increased by ~30% on average, and median half-life was extended from 9.45 hr to 11.2 hr (Figure 4A, right panel).

Of note, although TUT4/7 contribute to let-7 biogenesis, double knockdown of TUT4/7 (without simultaneous knockdown of TUT2) did not substantially affect the let-7 level (Heo et al., 2012). In fact, our transcriptome analyses show that mRNAs are globally upregulated, indicating that the changes in mRNA half-life observed in this study cannot be attributed to specific regulation of let-7 biogenesis.

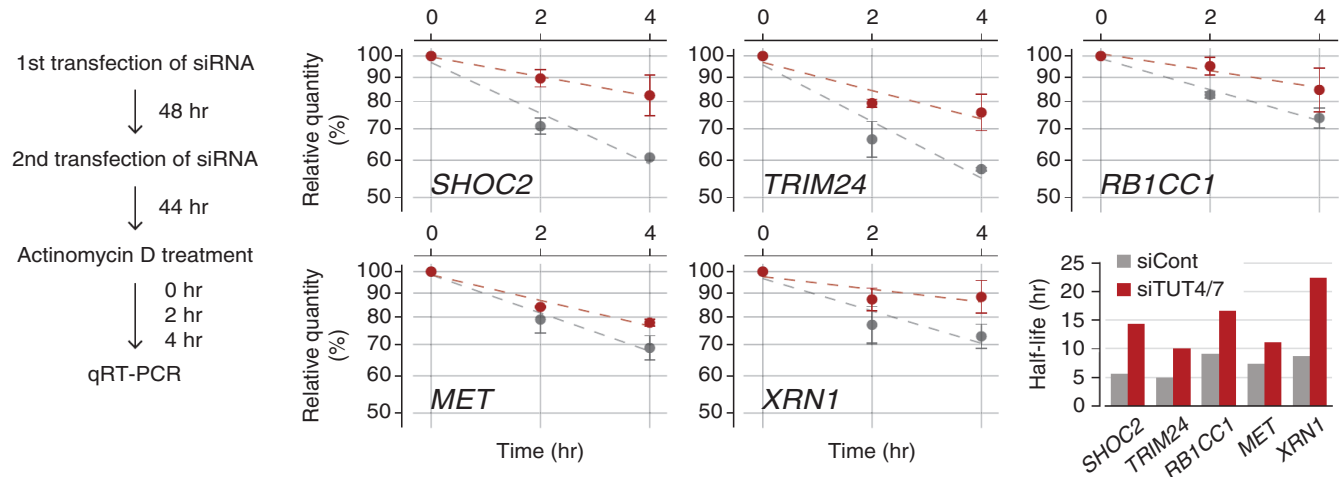
For validation of the impact of TUT4/7 depletion on mRNA stability, five mRNAs (*SHOC2*, *TRIM24*, *RB1CC1*, *MET*, and *XRN1*) were measured by quantitative RT-PCR after actinomycin D treatment (Figure 4B). None of these mRNAs contains a let-7 binding site with seed match in their 3' UTR, yet all of them showed increased stability when TUT4/7 were depleted. Therefore, our results demonstrate that TUT4/7 play an important role in bulk mRNA degradation in a let-7 independent manner.

Next, to examine the effect of overexpressed TUTase on mRNA expression, we carried out tethering experiments in HeLa cells (Figure 4C, left panel). A related experiment was reported recently in *Xenopus* oocytes: when *Xenopus* TUT7 homolog was tethered to the 3' UTR of luciferase reporter mRNA, luciferase activity was reduced without significant changes in mRNA, implicating translational repression (Lapointe and

A Measurement of mRNA half-life by RNA-seq



B Measurement of mRNA half-life by qRT-PCR



C Tethering experiment

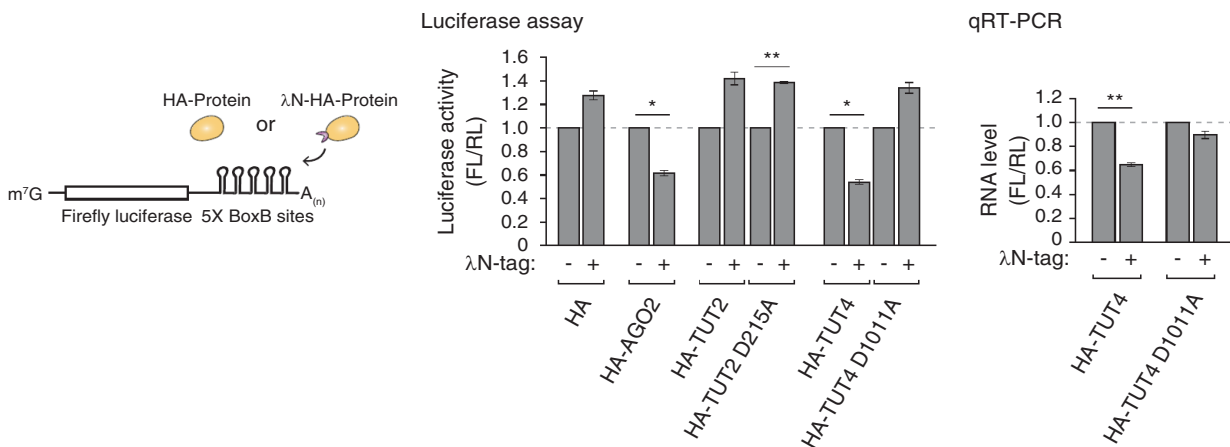


Figure 4. Uridylation Promotes mRNA Degradation

(A) Transcriptome-wide change of mRNA half-life determined by RNA-seq. Left: experimental scheme. HeLa cells were transfected twice and harvested at 0, 1, 2, and 4 hr following actinomycin D treatment. Center: changes of average mRNA half-life upon TUT4/7 knockdown from two biological replicates. The range of display is limited to between 0 and 30 hr for the better visual recognition (232 out of 1,829 mRNAs are outside of the view). The full list is available in Table S2. Right: distribution of mRNA half-lives in control or TUT4/7 knockdown cells. A box represents the first and third quartiles and an internal bar indicates median. Whiskers span between the ninth and the 91st percentiles. Half-lives of mRNAs are significantly extended by TUT4/7 knockdown (** $p = 4.06 \times 10^{-155}$, one-tailed paired Mann-Whitney U test). See Extended Experimental Procedures for the detailed description of procedure.

(legend continued on next page)

Wickens, 2013). However, because mRNA decay activity is generally suppressed in oocytes (Barckmann and Simonelig, 2013), it was unclear if the observation from frog oocytes can be generalized. For tethering experiments, we generated constructs that express proteins tagged with the λ N peptide that interacts with its specific binding sites (BoxB sites) in the 3' UTR of luciferase mRNA (Figures 4C and S4). Expression of λ N protein modestly increased luciferase expression nonspecifically for an unknown reason (Figure 4C, middle panel). Nevertheless, tethering of AGO2 repressed luciferase reporter expression (Figure 4C), as previously shown (Pillai et al., 2004), indicating that this is a valid system to test the effect of RNA silencing factors. Neither the negative control TUT2 nor its mutant repressed luciferase reporter expression. But when wild-type TUT4 was tethered to the reporter mRNA, luciferase activity was decreased to ~60% while such reduction was not observed with the catalytically dead point mutant (D1011A) of TUT4 (Figure 4C, middle panel), indicating that TUT4 suppressed gene expression via uridylation. Quantitative RT-PCR further showed that tethering of TUT4 induced a reduction of mRNA (Figure 4C, right panel). Thus, our results collectively indicate that TUT4/7 function as suppressors of gene expression through mRNA destabilization.

Uridylation Is Involved in miRNA-Induced Gene Silencing

Our model predicts that if a gene-specific inducer of deadenylation is introduced into cells, uridylation of the given transcript will take place, which in turn will facilitate RNA decay. To test our model, we examined the effect of miRNA as an example, which is well established to induce specific deadenylation of its complementary targets (Ameres and Zamore, 2013; Djuranovic et al., 2011; Huntzinger and Izaurralde, 2011; Krol et al., 2010).

We first analyzed the TAIL-seq data from our previous experiment where miR-1 mimic was transfected into HeLa cells (Chang et al., 2014b). As expected, miR-1 targets undergo deadenylation and subsequent downregulation following miR-1 mimic transfection (Figure 5A, middle and lower panels, respectively). Importantly, we detected a specific increase of uridylation on miR-1 targets whereas the rest of genes stayed largely unaffected (Figure 5A, upper panel). This result is consistent with our model that deadenylation leads to uridylation.

We next measured turnover rates of miR-1 targets with or without TUT4/7 knockdown. The mRNAs tested (*PTMA*, *ADAR*, and *PGM2*) are normally stable (half-lives >24 hr) in cells that do not contain miR-1 (Figure 5B, black line). When miR-1 was introduced, their half-lives were shortened to 6.5, 10.0, and 9.4 hr, respectively (Figure 5B, blue line). Upon TUT4/7 depletion, the miR-1 target mRNAs were stabilized (with extended half-lives of 14.0, 36.1, and 18.1 hr, respectively) (Figure 5B, red line). Therefore, TUT4/7 are necessary for the facilitated

decay of miRNA targets. We propose that other factors that cause deadenylation may also induce uridylation and decay, as shown here with an example of miR-1.

mRNA Decay Factors Remove Uridylated mRNAs

To understand downstream events of uridylation, we disrupted 5'–3' or 3'–5' exonucleolytic decay factors and examined the mRNA terminome (Figure S5A). The popsicle-shaped bars in Figure 6 display the relative quantity of reads with a U-tail (thick stem) or without a U-tail (thin stem). As U-tail frequencies vary depending on poly(A) tail length, different A-tail ranges are shown separately along the horizontal axis. For more information, the overall uridylation frequency and poly(A) length distribution are presented in Figures S5B and S5C, respectively.

In order to inhibit 5'–3' decay, we initially depleted a major 5'–3' exoribonuclease XRN1. Interestingly, interference of XRN1 resulted in a strong accumulation of uridylated mRNAs with short A-tails (≤ 25 nt) (Figure 6A). Additionally, when we depleted LSM1 (a component of the LSM1–7 complex that is known to facilitate decapping) or overexpressed dominant-negative mutants of the decapping complex (DCP1 and DCP2) (Chang et al., 2014a), we detected an increase of uridylation among short A-tailed mRNAs (≤ 25 nt) (Figure 6B). Short A-tailed mRNAs increased in abundance (particularly, in the 5–15 nt range) when the 5'–3' decay was suppressed. Note that the level of uridylated mRNA was upregulated substantially (U_1 – U_{3+}), accounting for the overall increase of mRNA reads in this range, while mRNAs without a U-tail did not change significantly (U_0). This result is consistent with a model that deadenylated, uridylated mRNAs are normally degraded rapidly by the 5'–3' decay factors while poly(A)⁺ mRNAs without U-tails are relatively stable. The LSM1–7 complex is known to preferentially bind to RNAs with 3' terminal uridyl residues (Chowdhury et al., 2007; Sharif and Conti, 2013; Song and Kiledjian, 2007; Zhou et al., 2014) and facilitate decapping through PATL1 (Pat1p in yeast) (Marnef and Standart, 2010; Wilusz and Wilusz, 2013). Thus, a short U-tail may first be recognized by the LSM1–7 complex which in turn facilitates decapping (by the DCP1/2 complex) and subsequent 5'–3' degradation (by XRN1).

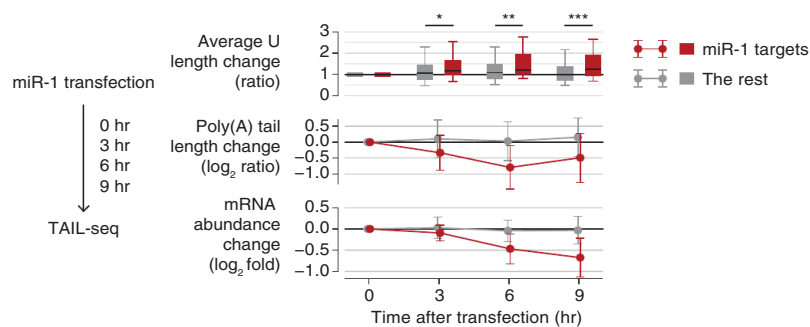
We also investigated the contribution of the 3'–5' decay pathway by depleting 3' exonucleolytic factors. When we knocked-down RRP41, a core subunit of human exosome, we detected a substantial accumulation of uridylated mRNAs with short A-tails (Figure 6C). Combinatorial knockdown of RRP41 and XRN1 resulted in a more pronounced increase of uridylation (Figure 6C). Therefore, both decay pathways (5'–3' and 3'–5') may act at the downstream of uridylation. We also tested a 3'–5' exonuclease DIS3L2 which is related to DIS3 and DIS3L. While DIS3 and DIS3L function as components of exosome,

(B) Measurement of mRNA half-life by qRT-PCR. Left: the experimental scheme. Right: following 0, 2, and 4 hr of actinomycin D treatment, relative abundance (y axis) of five selected genes were measured. For normalization, *GAPDH* mRNA was used because it was highly stable (half-life > 24 hr, data not shown) and did not change noticeably by TUT4/7 depletion. Error bar represents SEM ($n = 3$). Half-lives are calculated by linear fitting of the log-transformed exponential decay function.

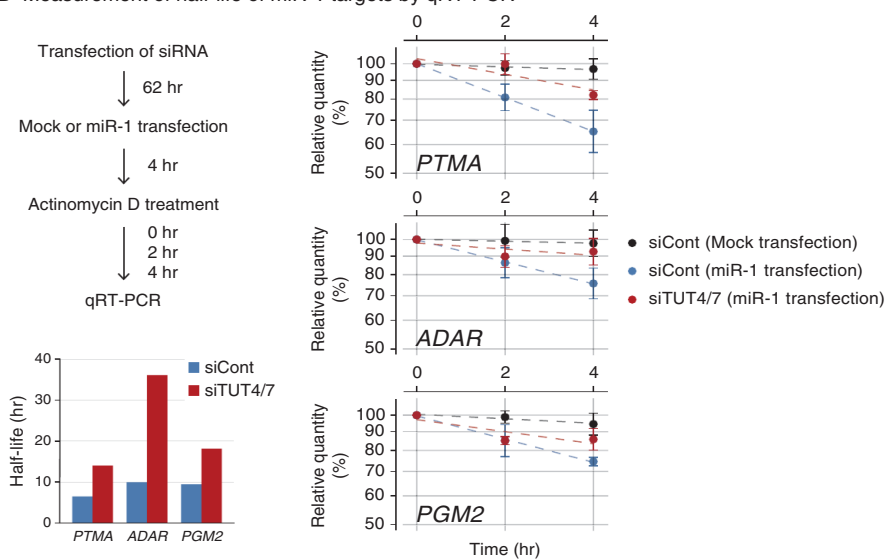
(C) Left: schematic representation of reporter assay system with the λ N tethering. Center: reporter (firefly) luciferase activity was measured and normalized to Renilla luciferase activity ($n = 3$). Right: reporter mRNA levels were determined by qRT-PCR ($n = 4$). Error bars represent SEM. Luciferase activity or RNA level were significantly reduced when AGO2 or TUT4 were tethered (* $p < 0.01$, ** $p < 0.001$; two-tailed t test).

See also Figure S4.

A Changes in uridylation after miR-1 transfection



B Measurement of half-life of miR-1 targets by qRT-PCR



abundance (y axis) of miR-1 target mRNAs were measured. For the normalization, highly stable *GAPDH* mRNA was used because it did not change significantly by siTUT4/7 or miR-1 transfection. Error bar represents SEM (n = 3). Half-lives are determined by linear fitting of the log-transformed exponential decay function.

DIS3L2 is known to work independently from exosome (Lubas et al., 2013; Malecki et al., 2013). It was recently shown that DIS3L2 preferentially binds to long U-tails of pre-let-7 and is involved in turnover of pre-let-7 and some mRNAs in yeast and human (Chang et al., 2013; Faehnle et al., 2014; Lubas et al., 2013; Malecki et al., 2013; Ustianenko et al., 2013). Our TAIL-seq experiment shows that DIS3L2 depletion results in a modest accumulation of uridylated reads (Figure 6C). Thus, although we cannot rule out the possibility of indirect effects, our results suggest that multiple decay pathways may participate in the removal of uridylated mRNAs. Due to the technical limitation of knock-down experiment, it is currently unclear which pathway plays a dominant role.

Interestingly, mRNAs with an oligo-U-tail (U₂ and U₃₊) responded more sensitively to the suppression of decay factors than those with a mono-U-tail (U₁), suggesting that oligo-uridylated mRNAs are more rapidly degraded than mono-uridylated mRNAs (Figures 6A–6C). Taken together, we propose that

Figure 5. Uridylation Facilitates miRNA-Mediated mRNA Decay

(A) Changes in uridylation after miR-1 transfection. Left: experimental scheme. miR-1 was transfected into HeLa cells and the cells were harvested after the indicated time for TAIL-seq. Targets are the transcripts with ≥ 1 miR-1 3' UTR site and down-regulated by $\geq 30\%$ on 12 hr posttransfection of miR-1 (Guo et al., 2010). Right top: average U length change relative to 0 hr is shown in each time point. Average U length per tail is the number of uridines on short A-tails (5–25 nt) divided by the total number of reads with short A-tails. Box represents the interval between the first and third quartiles, and the internal bar indicates the median. Whiskers span between the ninth and 91st percentiles. Average U length of miR-1 target is significantly extended after miR-1 transfection (* $p = 0.0152$, ** $p = 0.00318$, *** $p = 5.79 \times 10^{-4}$; one-tailed Mann-Whitney U test). Right middle: poly(A) tail length change relative to 0 hr. The length change is represented by log₂ odds ratio between long tails (> 25 nt) and short tails (≤ 25 nt) in one among 3, 6, or 9 hr and 0 hr. A negative value (< 0) indicates increase of the fraction of short tails compared to 0 hr. Error bars indicate SD among mRNAs. The portion of short poly(A) tails expanded more for miR-1 targets than the others ($p = 1.80 \times 10^{-6}$ for 3 hr, $p = 8.47 \times 10^{-13}$ for 6 hr, $p = 1.48 \times 10^{-11}$ for 9 hr; one-tailed Mann-Whitney U test). Right bottom: mRNA abundance (poly(A)⁺ tag counts) change relative to 0 hr. Error bars indicate SD among mRNAs. Expression levels of miR-1 targets were decreased more than the rest transcripts ($p = 2.09 \times 10^{-4}$ for 3 hr, $p = 2.65 \times 10^{-14}$ for 6 hr, $p = 5.46 \times 10^{-18}$ for 9 hr; one-tailed Mann-Whitney U test).

(B) Measurement of half-life of miR-1 targets by qRT-PCR. Left: the experimental scheme. Following siRNA transfection for 62 hr, HeLa cells were transfected with miR-1 or mock transfected. After 4 hr, actinomycin D was treated and cells were harvested at 0, 1, 2, and 4 hr. Right: relative

oligo-uridylated mRNAs are subject to degradation by multiple factors, and an oligo-U-tail may serve as a decay mark for nonfunctional, deadenylated mRNAs.

DISCUSSION

In conclusion, this study reveals an integral and general role of oligo-uridylation in mammalian mRNA decay (model shown in Figure 7). Upon deadenylation, mRNAs (with A-tails shorter than ~ 25 nt) lose PABP and instead gain a U-tail by the redundant action of TUT4 and TUT7. The oligo-U-tail triggers decay by serving as a mark that is recognized by downstream decay factors. Thus, TUT4/7 function as the “writers” of the decay mark. It will be interesting in the future to identify the “readers” of the oligo-U-tail and to ask if this modification can be reversed by “erasers.” The LSM1–7 complex and DIS3L2 are likely candidates that recognize the oligo-U marks, but further investigations will be necessary to understand which factor(s) recognize the

RNAi followed by TAIL-seq

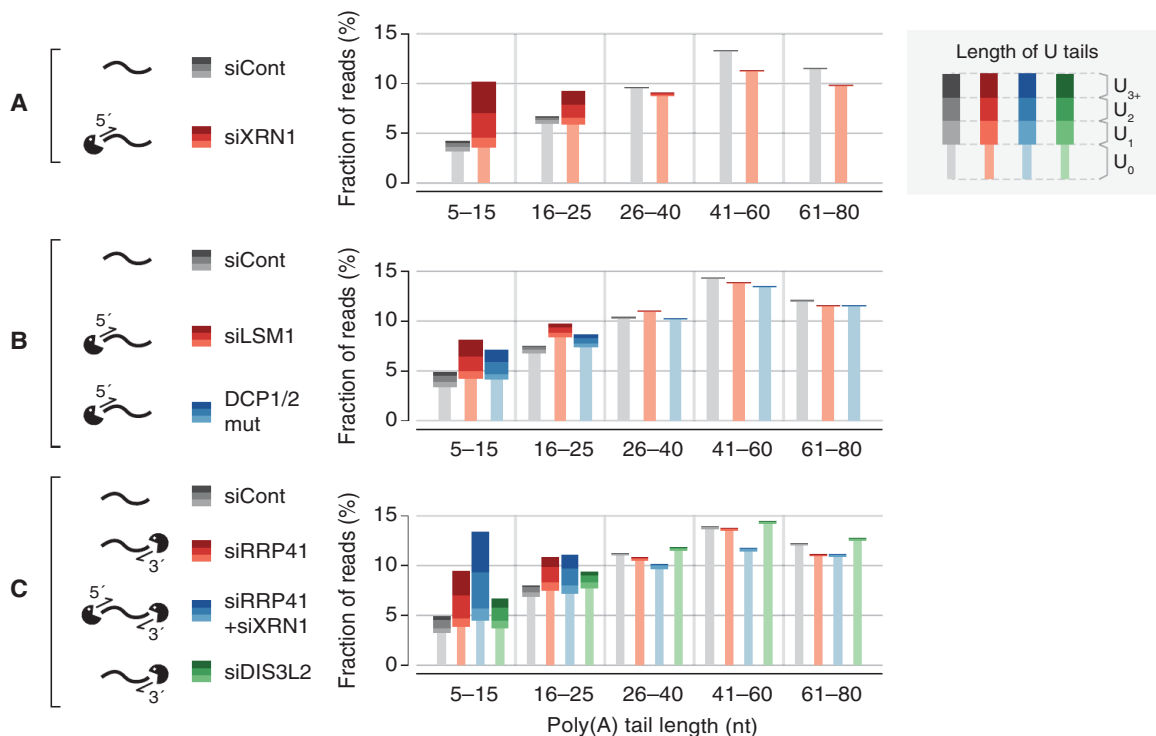


Figure 6. The 5' and 3' mRNA Decay Factors Degrade Uridylated mRNAs

(A–C) Changes of poly(A) tail and uridylation upon knockdown of decay factor(s) detected by small-scale TAIL-seq (with Illumina MiSeq). Fraction of mRNA reads out of the total poly(A)⁺ mRNA reads is shown in each poly(A) tail size range. Narrow bars represent reads without U-tails (U₀) and wider bars indicate uridylated reads (U₁–U₃₊). The “DCP1/2 mut” sample derived from cells coexpressed of dominant-negative mutants of DCP1 and DCP2 (DCP1a-GSSG and DCP2-E148Q, respectively).

See also Figure S5.

oligo-U-tails mainly, whether there is any additional factor(s) that binds to the oligo-U-tails, and what is the molecular basis of the specific recognition (Lee et al., 2014).

It is intriguing that TUT4/7 are capable of measuring poly(A) length (Figure 2). Poly(A) tail is unlikely to form a certain structure through base-pairing, so we do not yet understand how RNA with a poly(A) tail is discriminated by TUT4/7. It would be interesting to carry out structural studies on TUT4/7 and RNA with an A-tail of various length. Furthermore, we found that PABPC1 preferentially protects long poly(A) tails from uridylation (Figure 3). This specific inhibitory effect may come from the length-dependent binding of PABPC1 (Kühn and Pieler, 1996; Sachs et al., 1987). Thus, the combined action of TUT4/7 and PABP may selectively mark nonfunctional mRNAs while translationally active polyadenylated mRNAs are refractory to uridylation. Consequently, TUT4/7-mediated uridylation may provide the molecular basis for the tight control of mRNA stability according to poly(A) tail length.

We observed that oligo-uridylated mRNAs (with ≥ 2 uridines) are more sensitive than mono-uridylated mRNAs to the knockdown of TUT4/7 and decay factors. Moreover, oligo-U-tails are found in a narrow range of short A-tail length while mono-U-tails are more loosely distributed and found in polyadenylated

mRNAs as well to some extent (Chang et al., 2014b). Thus, mono-uridylation appears to be less specific than oligo-uridylation and may be catalyzed in part by a TUT(s) other than TUT4/7. Furthermore, mono-U-tails may be too short to recruit decay factors effectively. Oligo-uridylated mRNAs are detected more frequently after depletion of decay factors, indicating that they are less stable than mono-uridylated mRNAs in control cells. Therefore, oligo-U-tails are likely to have a stronger effect in decay than mono-U-tails do. In fission yeast and plants, it is currently unclear if there is such a distinction between oligo-U-tails and mono-U-tails because only a small number of reads from cloning has been analyzed thus far.

Our transcriptome-wide analyses allowed us to propose a general model for the decay of poly(A)⁺ mRNAs. In addition, given that poly(A)[−] histone mRNA was also proposed to be uridylated by TUT4 (Schmidt et al., 2011; Su et al., 2013), it is possible that both poly(A)⁺ mRNAs and poly(A)[−] mRNAs are degraded by the same general principle involving uridylation although there may be some differences in details such as the choice of downstream decay factors. In fact, we detected uridylation on histone mRNAs and on trimmed decay intermediates lacking poly(A) tail and these U-tails were also dependent on TUT4/7 (data not shown).

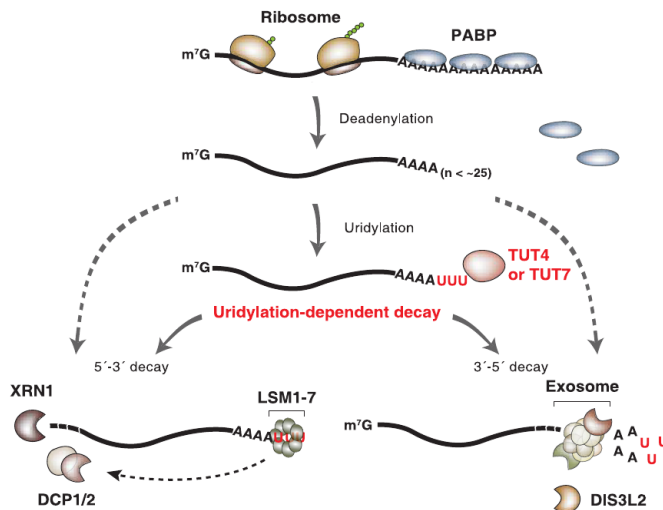


Figure 7. Model for Uridylation-Dependent mRNA Decay in Humans
mRNA decay is generally initiated by deadenylation. PABP proteins are dissociated from mRNA as poly(A) tail becomes shorter (less than ~25 nt). TUT4 and TUT7 act redundantly to uridylate mRNAs with a short A-tail. The U-tail is in turn recognized by the downstream decay factors (uridylation-dependent mRNA decay pathway). The LSM1–7 complex binds to the U-tail and facilitates decapping by the DCP1/2 complex. Decapped mRNAs are degraded by the 5'–3' exonuclease XRN1. Alternatively, the U-tail is recognized by exosome or DIS3L2 that degrade mRNA exonucleolytically from the 3' end. It is currently unclear if and what fraction of deadenylated mRNAs are degraded through uridylation-independent alternative pathways (indicated with gray dashed lines).

In addition, we found that miR-1 transfection results in an increased uridylation and facilitated decay of its targets (Figure 5). These results suggest that uridylation contributes to miRNA-mediated gene silencing by removing the body of deadenylated mRNAs. Uridylation may be involved in other decay and surveillance pathways in mammals, playing a general role. It is noted that we cannot currently assess if and to what extent uridylation-independent alternative pathway(s) contribute to bulk mRNA decay.

Tailing of mRNA is found in many eukaryotes, with some notable differences among the species. In filamentous fungus *Aspergillus nidulans*, mRNAs carry 3' tails mixed with cytidine and uridine (Morozov et al., 2010). In a double deletion mutant of noncanonical PAPs, CutA and CutB, this “CUCU” modification was abrogated, and transcripts were stabilized, indicating that a CUCU tail also serves as a decay mark despite the difference in base composition (Morozov et al., 2010, 2012). In plants, although uridylation occurs similarly to mammals, mRNA half-life did not change in the *urt1* mutants, and the reason underlying the difference is currently unclear (Sement et al., 2013). Another variation among the species is that uridylation occurs selectively on deadenylated mRNAs in mammals and plants whereas uridylation appears to be independent of poly(A) tail length in *S. pombe* and *A. nidulans* (Morozov et al., 2010; Rissland and Norbury, 2009). Deadenylation may not be a prerequisite for uridylation in fungi as they possess shorter poly(A) tail (20–30 nt in median) than mammals (60–100 nt in median) and plants (50–60 nt in median) (Chang et al., 2014b; Morozov et al., 2010; Subtelny et al., 2014). Thus, further

investigations are clearly necessary to delineate the commonalities and differences of uridylation in diverse systems.

Tailing-mediated decay is deeply conserved and found even in prokaryotes where mRNAs typically end with stem loop structure and are degraded in an adenylation-dependent manner (Belasco, 2010; Houseley et al., 2006). An oligo-A-tail serves as a single-stranded toehold for 3' exonucleases that are otherwise hindered by the terminal stem loop. A related phenomenon was observed in budding yeast where noncanonical PAPs, Trf4 and Trf5, adenylate defective nuclear RNAs and facilitate their degradation by exosome (Houseley et al., 2006; Norbury, 2013). Our current work shows that mammalian cytoplasmic mRNAs use uridylation, instead of adenylation, to promote mRNA decay. Together with previous findings (Morozov et al., 2010; Mullen and Marzluff, 2008; Rissland and Norbury, 2009), our study establishes a fundamental and conserved role for tailing in the mRNA decay pathways.

EXPERIMENTAL PROCEDURES

Construction of TAIL-Seq Library

TAIL-seq was carried out as described previously (Chang et al., 2014b). Briefly, 25–50 µg of total RNA was extracted using TRIzol (Invitrogen), purified with RNeasy MinElute column (QIAGEN), and rRNA-depleted by using Ribo-Zero kit (Epicentre). The RNAs were ligated to the biotinylated 3' adaptor and partially digested by RNase T1 (Ambion). The fragmented RNAs were precipitated with streptavidin beads, phosphorylated at the 5' end, and gel purified (500–1,000 nt). The purified RNAs were ligated to the 5' adaptor, reverse-transcribed, and amplified by PCR. The cDNA libraries were mixed with PhiX control library v3 (Illumina) and spike-in mixture and then sequenced by paired-end run (51 × 251 cycles) on Illumina MiSeq (small-scale TAIL-seq) or HiSeq 2500. Resulting data were processed as previously described (Chang et al., 2014b). See also Extended Experimental Procedures.

In Vitro Uridylation Assay

For immunoprecipitation of FLAG-TUTases, HEK293T cells grown on 10 cm dishes were collected 48 hr after transfection with FLAG-TUTase expression plasmids (full-length human TUT4 [1–1,640 aa] and human TUT7 [1–1,495 aa]). The cells were incubated in ice-cold Buffer D (200 mM KCl, 10 mM Tris-HCl [pH 8.0], 0.2 mM EDTA) for 20 min followed by sonication on ice and centrifugation twice for 15 min at 4°C. The supernatant was incubated with 5 µl of anti-FLAG antibody-conjugated agarose beads (anti-FLAG M2 affinity gel, Sigma) with constant rotation for 1 hr at 4°C. The beads were washed six times with Buffer D. Uridylation reaction was done in a total volume of 30 µl in 3.2 mM MgCl₂, 1 mM DTT, 0.67 U/µl RNase inhibitor (Promega, N2515), 0.25 mM UTP, 0.45 nM of 5' end-labeled RNA, and 15 µl of immunopurified proteins on beads or 3X Flag-peptide (Sigma) eluted proteins in Buffer D. When uridylation assay was done with recombinant TUT7 (951–1,495 aa), 14 nM of protein was used. The reaction mixture was incubated at 37°C for up to 10 min. For uridylation assay in the presence of PABPC1, 10–40 nM of recombinant human PABPC1 (Origene, TP307354) was preincubated with RNA for 10 min and then uridylation was carried out by adding 160 nM of recombinant TUT7 (951–1,495 aa). Buffer D with final 300 mM KCl was used when uridylation assay was carried out in the presence of PABPC1. The RNA was purified from the reaction mixture by phenol extraction and run on 6% polyacrylamide sequencing gel with 7 M urea (20 × 40 cm, 0.4 mm thick) at constant 1,500 V for 2 hr. The gel was exposed to phosphor imaging plate (Fujifilm) and read by Typhoon FLA 7000 (GE Healthcare). The signal intensity profile was quantified using MultiGauge v3.0 (Fujifilm). In Figure S2D, 12.5% polyacrylamide gel was used. The *SHOC2* 3' UTR and *CALM1* 3' UTR were selected as RNA substrates as they do not contain homopolymeric adenosines at the 3' end. RNAs were synthesized by ST Pharm.

The list of RNA oligos is shown in Table S3.

ACCESSION NUMBERS

The NCBI Gene Expression Omnibus (GEO) accession number for the sequenced reads reported in this paper is GSE59628.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, five figures, and three tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2014.10.055>.

AUTHOR CONTRIBUTIONS

J.L., M.H., H.C., and V.N.K. designed experiments. J.L., M.H., and S.C.K. performed biochemical and cell biological experiments. H.C. carried out computational analyses. D.K.S. and D.J.P. prepared recombinant proteins. J.L., M.H., H.C., and V.N.K. wrote the manuscript.

ACKNOWLEDGMENTS

We are grateful to members of our laboratory for discussion and technical help, especially Eunji Kim for help with plasmid cloning. We thank Dr. E. Izaurralde for insightful suggestions and the gifts of DCP1/2 mutant plasmids and Dr. G. Dreyfuss for the gift of anti-PABPC1 antibody. This work was supported by IBS-R008-D1 of Institute for Basic Science from the Ministry of Science, ICT and Future Planning of Korea (J.L., M.H., H.C., S.C.K., and V.N.K.) and the BK21 Research Fellowships from the Ministry of Education of Korea (J.L.).

Received: July 16, 2014

Revised: September 24, 2014

Accepted: October 20, 2014

Published: December 4, 2014

REFERENCES

- Ameres, S.L., and Zamore, P.D. (2013). Diversifying microRNA sequence and function. *Nat. Rev. Mol. Cell Biol.* **14**, 475–488.
- Aravind, L., and Koonin, E.V. (1999). DNA polymerase beta-like nucleotidyl-transferase superfamily: identification of three new families, classification and evolutionary history. *Nucleic Acids Res.* **27**, 1609–1618.
- Baer, B.W., and Kornberg, R.D. (1983). The protein responsible for the repeating structure of cytoplasmic poly(A)-ribonucleoprotein. *J. Cell Biol.* **96**, 717–721.
- Barckmann, B., and Simonelig, M. (2013). Control of maternal mRNA stability in germ cells and early embryos. *Biochim. Biophys. Acta* **1829**, 714–724.
- Belasco, J.G. (2010). All things must pass: contrasts and commonalities in eukaryotic and bacterial mRNA decay. *Nat. Rev. Mol. Cell Biol.* **11**, 467–478.
- Chang, H.M., Triboulet, R., Thornton, J.E., and Gregory, R.I. (2013). A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway. *Nature* **497**, 244–248.
- Chang, C.T., Bercovich, N., Loh, B., Jonas, S., and Izaurralde, E. (2014a). The activation of the decapping enzyme DCP2 by DCP1 occurs on the EDC4 scaffold and involves a conserved loop in DCP1. *Nucleic Acids Res.* **42**, 5217–5233.
- Chang, H., Lim, J., Ha, M., and Kim, V.N. (2014b). TAIL-seq: genome-wide determination of poly(A) tail length and 3' end modifications. *Mol. Cell* **53**, 1044–1052.
- Chowdhury, A., Mukhopadhyay, J., and Tharun, S. (2007). The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs. *RNA* **13**, 998–1016.
- Djuranovic, S., Nahvi, A., and Green, R. (2011). A parsimonious model for gene regulation by miRNAs. *Science* **331**, 550–553.
- Dreyfus, M., and Régnier, P. (2002). The poly(A) tail of mRNAs: bodyguard in eukaryotes, scavenger in bacteria. *Cell* **111**, 611–613.
- Eliseeva, I.A., Lyabin, D.N., and Ovchinnikov, L.P. (2013). Poly(A)-binding proteins: structure, domain organization, and activity regulation. *Biochemistry Mosc.* **78**, 1377–1391.
- Faehnle, C.R., Walleshauser, J., and Joshua-Tor, L. (2014). Mechanism of Dis3l2 substrate recognition in the Lin28-let-7 pathway. *Nature* **514**, 252–256.
- Garneau, N.L., Wilusz, J., and Wilusz, C.J. (2007). The highways and byways of mRNA decay. *Nat. Rev. Mol. Cell Biol.* **8**, 113–126.
- Guo, H., Ingolia, N.T., Weissman, J.S., and Bartel, D.P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* **466**, 835–840.
- Heo, I., Ha, M., Lim, J., Yoon, M.J., Park, J.E., Kwon, S.C., Chang, H., and Kim, V.N. (2012). Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. *Cell* **151**, 521–532.
- Hoefig, K.P., Rath, N., Heinz, G.A., Wolf, C., Dameris, J., Schepers, A., Kremer, E., Ansel, K.M., and Heissmeyer, V. (2013). Eri1 degrades the stem-loop of oligouridylated histone mRNAs to induce replication-dependent decay. *Nat. Struct. Mol. Biol.* **20**, 73–81.
- Houseley, J., and Tollervey, D. (2009). The many pathways of RNA degradation. *Cell* **136**, 763–776.
- Houseley, J., LaCava, J., and Tollervey, D. (2006). RNA-quality control by the exosome. *Nat. Rev. Mol. Cell Biol.* **7**, 529–539.
- Huntzinger, E., and Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat. Rev. Genet.* **12**, 99–110.
- Khanam, T., Muddashetty, R.S., Kahvejian, A., Sonenberg, N., and Brosius, J. (2006). Poly(A)-binding protein binds to A-rich sequences via RNA-binding domains 1+2 and 3+4. *RNA Biol.* **3**, 170–177.
- Krol, J., Loedige, I., and Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.* **11**, 597–610.
- Kühn, U., and Pieler, T. (1996). Xenopus poly(A) binding protein: functional domains in RNA binding and protein-protein interaction. *J. Mol. Biol.* **256**, 20–30.
- Lapointe, C.P., and Wickens, M. (2013). The nucleic acid-binding domain and translational repression activity of a Xenopus terminal uridylyl transferase. *J. Biol. Chem.* **288**, 20723–20733.
- Lee, M., Kim, B., and Kim, V.N. (2014). Emerging roles of RNA modification: m(6)A and U-tail. *Cell* **158**, 980–987.
- Liu, X., Zheng, Q., Vrettos, N., Maragkakis, M., Alexiou, P., Gregory, B.D., and Mourelatos, Z. (2014). A MicroRNA precursor surveillance system in quality control of MicroRNA synthesis. *Mol. Cell* **55**, 868–879.
- Lubas, M., Damgaard, C.K., Tomecki, R., Cysewski, D., Jensen, T.H., and Dziembowski, A. (2013). Exonuclease hDIS3L2 specifies an exosome-independent 3'-5' degradation pathway of human cytoplasmic mRNA. *EMBO J.* **32**, 1855–1868.
- Malecki, M., Viegas, S.C., Carneiro, T., Golik, P., Dressaire, C., Ferreira, M.G., and Arraiano, C.M. (2013). The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. *EMBO J.* **32**, 1842–1854.
- Marnef, A., and Standart, N. (2010). Pat1 proteins: a life in translation, translation repression and mRNA decay. *Biochem. Soc. Trans.* **38**, 1602–1607.
- Martin, G., and Keller, W. (2007). RNA-specific ribonucleotidyl transferases. *RNA* **13**, 1834–1849.
- Morozov, I.Y., Jones, M.G., Razak, A.A., Rigden, D.J., and Caddick, M.X. (2010). CUCU modification of mRNA promotes decapping and transcript degradation in *Aspergillus nidulans*. *Mol. Cell Biol.* **30**, 460–469.
- Morozov, I.Y., Jones, M.G., Gould, P.D., Crome, V., Wilson, J.B., Hall, A.J., Rigden, D.J., and Caddick, M.X. (2012). mRNA 3' tagging is induced by nonsense-mediated decay and promotes ribosome dissociation. *Mol. Cell Biol.* **32**, 2585–2595.
- Mullen, T.E., and Marzluff, W.F. (2008). Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. *Genes Dev.* **22**, 50–65.

- Norbury, C.J. (2013). Cytoplasmic RNA: a case of the tail wagging the dog. *Nat. Rev. Mol. Cell Biol.* **14**, 643–653.
- Parker, R., and Song, H. (2004). The enzymes and control of eukaryotic mRNA turnover. *Nat. Struct. Mol. Biol.* **11**, 121–127.
- Pillai, R.S., Artus, C.G., and Filipowicz, W. (2004). Tethering of human Ago proteins to mRNA mimics the miRNA-mediated repression of protein synthesis. *RNA* **10**, 1518–1525.
- Rissland, O.S., and Norbury, C.J. (2009). Decapping is preceded by 3' uridylation in a novel pathway of bulk mRNA turnover. *Nat. Struct. Mol. Biol.* **16**, 616–623.
- Rissland, O.S., Mikulasova, A., and Norbury, C.J. (2007). Efficient RNA polyuridylation by noncanonical poly(A) polymerases. *Mol. Cell. Biol.* **27**, 3612–3624.
- Sachs, A.B., Davis, R.W., and Kornberg, R.D. (1987). A single domain of yeast poly(A)-binding protein is necessary and sufficient for RNA binding and cell viability. *Mol. Cell. Biol.* **7**, 3268–3276.
- Schmidt, M.J., West, S., and Norbury, C.J. (2011). The human cytoplasmic RNA terminal U-transferase ZCCHC11 targets histone mRNAs for degradation. *RNA* **17**, 39–44.
- Sement, F.M., Ferrier, E., Zuber, H., Merret, R., Alioua, M., Deragon, J.M., Bousquet-Antonelli, C., Lange, H., and Gagliardi, D. (2013). Uridylation prevents 3' trimming of oligoadenylated mRNAs. *Nucleic Acids Res.* **41**, 7115–7127.
- Sharif, H., and Conti, E. (2013). Architecture of the Lsm1-7-Pat1 complex: a conserved assembly in eukaryotic mRNA turnover. *Cell Reports* **5**, 283–291.
- Shen, B., and Goodman, H.M. (2004). Uridine addition after microRNA-directed cleavage. *Science* **306**, 997.
- Slevin, M.K., Meaux, S., Welch, J.D., Bigler, R., Miliani de Marval, P.L., Su, W., Rhoads, R.E., Prins, J.F., and Marzluff, W.F. (2014). Deep sequencing shows multiple oligouridylations are required for 3' to 5' degradation of histone mRNAs on polyribosomes. *Mol. Cell* **53**, 1020–1030.
- Song, M.G., and Kiledjian, M. (2007). 3' Terminal oligo U-tract-mediated stimulation of decapping. *RNA* **13**, 2356–2365.
- Su, W., Slepnev, S.V., Slevin, M.K., Lyons, S.M., Ziemniak, M., Kowalska, J., Darzynkiewicz, E., Jemielity, J., Marzluff, W.F., and Rhoads, R.E. (2013). mRNAs containing the histone 3' stem-loop are degraded primarily by decapping mediated by oligouridylation of the 3' end. *RNA* **19**, 1–16.
- Subtelny, A.O., Eichhorn, S.W., Chen, G.R., Sive, H., and Bartel, D.P. (2014). Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature* **508**, 66–71.
- Thornton, J.E., Chang, H.M., Piskounova, E., and Gregory, R.I. (2012). Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). *RNA* **18**, 1875–1885.
- Ustianenko, D., Hrossova, D., Potesil, D., Chalupnikova, K., Hrazdilova, K., Pachernik, J., Cetkovska, K., Uldrijan, S., Zdrahal, Z., and Vanacova, S. (2013). Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. *RNA* **19**, 1632–1638.
- Wilusz, C.J., and Wilusz, J. (2013). Lsm proteins and Hfq: Life at the 3' end. *RNA Biol.* **10**, 592–601.
- Zhou, L., Zhou, Y., Hang, J., Wan, R., Lu, G., Yan, C., and Shi, Y. (2014). Crystal structure and biochemical analysis of the heptameric Lsm1-7 complex. *Cell Res.* **24**, 497–500.

Subnucleosomal Structures and Nucleosome Asymmetry across a Genome

Ho Sung Rhee,^{1,2} Alain R. Bataille,¹ Liye Zhang,^{1,3} and B. Franklin Pugh^{1,*}

¹Center for Eukaryotic Gene Regulation, Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802, USA

²Present address: Departments of Pathology and Cell Biology, Neurology, and Neuroscience, Center for Motor Neuron Biology and Disease, Columbia University Medical Center, New York, NY 10032, USA

³Present address: Department of Medicine, Computational Biomedicine, Boston University School of Medicine, 72 East Concord Street, E-648, Boston, MA 02118, USA

*Correspondence: bfp2@psu.edu

<http://dx.doi.org/10.1016/j.cell.2014.10.054>

SUMMARY

Genes are packaged into nucleosomal arrays, each nucleosome typically having two copies of histones H2A, H2B, H3, and H4. Histones have distinct post-translational modifications, variant isoforms, and dynamics. Whether each histone copy within a nucleosome has distinct properties, particularly in relation to the direction of transcription, is unknown. Here we use chromatin immunoprecipitation-exonuclease (ChIP-exo) to resolve the organization of individual histones on a genomic scale. We detect widespread subnucleosomal structures in dynamic chromatin, including what appear to be half-nucleosomes consisting of one copy of each histone. We also detect interactions of H3 tails with linker DNA between nucleosomes, which may be negatively regulated by methylation of H3K36. Histone variant H2A.Z is enriched on the promoter-distal half of the +1 nucleosome, whereas H2BK123 ubiquitylation and H3K9 acetylation are enriched on the promoter-proximal half in a transcription-linked manner. Subnucleosome asymmetries might serve as molecular beacons that guide transcription.

INTRODUCTION

Nearly every gene in a eukaryotic nucleus is packaged into chromatin by an array of nucleosomes (Jiang and Pugh, 2009; Rando and Ahmad, 2007; Segal and Widom, 2009). How these genic arrays are structured in relation to transcription is only partly understood. The first nucleosome in each array typically resides at a canonical distance from the transcription start site (TSS) and at the edge of a 5' nucleosome-free promoter region (NFR). A subset of quiescent genes, typically regulated by the SAGA complex, may have nucleosomes over their promoters that are lost or depleted upon gene activation. The +1 nucleosome is the gateway to transcription as it is the first nucleosome encountered by the transcription machinery. Arrays continue into gene bodies, having nucleosome repeat lengths (NRLs) of ~165 bp

in budding yeast. This regularity dissipates toward the middle of genes.

In the wake of DNA replication, nucleosomes are assembled by chaperones. They first escort dimers of histones H3/H4 into tetramer intermediates on ~60 bp of DNA; these tetramer intermediates are then rapidly flanked by two sets of H2A/H2B dimers that together wrap ~147 bp of DNA ~1.65 times around the octamer core (Luger et al., 2012). This basic two-step assembly process has been a tenet in chromatin biology for over 25 years (Kornberg and Lorch, 1999). However, early studies hinted at alternative pathways that produce subnucleosomal particles (Weintraub et al., 1975; Weintraub et al., 1976). Regardless, nucleosomal and subnucleosomal structures with respect to the organization of individual histones and their chromatin context have not been defined on a genome-wide scale.

A nucleosome has two-fold symmetry of histone organization and thus might have a symmetrical distribution of histones, variants, and modifications about its dyad axis. However, RNA polymerase II engages the NFR-proximal face of a nucleosome differently than it engages its distal back-end, as it transcribes a gene. Either as a cause or a consequence of an asymmetric polymerase-nucleosome relationship, the levels of histones, variants, and their modifications might be asymmetrically distributed within specific nucleosomes.

The chromatin immunoprecipitation-exonuclease (ChIP-exo) assay locates formaldehyde-induced protein-DNA crosslinks along a genome at very high resolution in vivo (Rhee and Pugh, 2011, 2012b). Here we apply ChIP-exo to the budding yeast four core histones, histone variant H2A.Z, the linker histone H1, and the transcription-linked histone modifications H3K4me3, H3K36me3, H3K79me2/3, H3K9ac, and H2BK123ub. With this, we examine nucleosome substructure and symmetry at the ~60,000 nucleosome-occupied sites in the budding yeast genome. Our study suggests a surprising model of intra- and internucleosomal histone interplay that may reflect a variety of subnucleosomal structures and their dynamics. A substantial fraction of all genes display differential histone occupancy on one-half of their nucleosomes versus the other. Nucleosomes at specific positions within arrays have an asymmetric organization of transcription-linked histone variants and modifications. Together, these findings paint a

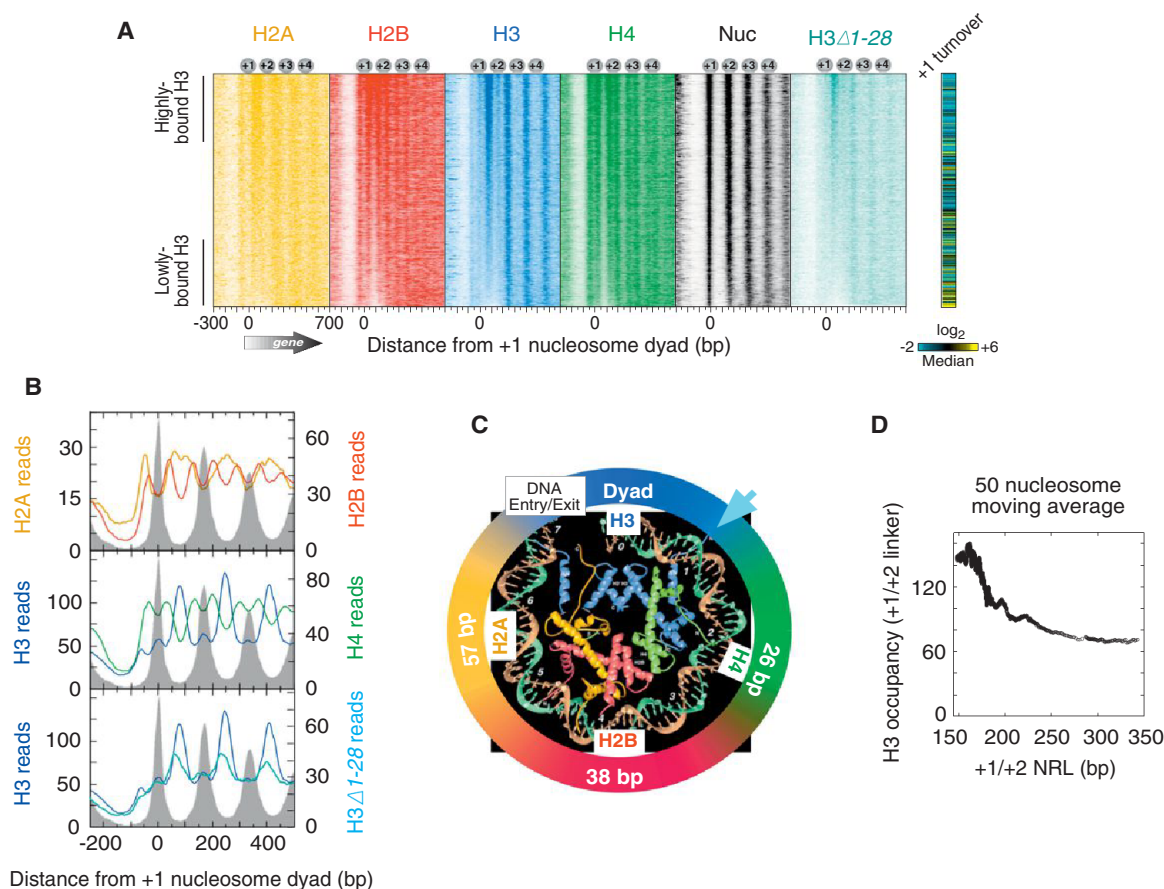


Figure 1. Subnucleosomal Detection of Histones across Yeast Genes

(A) Histone occupancy (color intensity) as detected by ChIP-exo was plotted relative to +1 nucleosome dyads ("Nuc," defined by MNase) at annotated mRNA genes (4,738 rows; values reported in Table S3). Rows were sorted by H3 occupancy in the linker between the +1 and +2 nucleosomes (Table S2). Histone turnover rate in the +1 nucleosome region is from Dion et al. (2007), where yellow, black, and blue represent high, medium, and low turnover, respectively.

(B) Composite (average) plot of (A). Gray fill indicates nucleosome midpoint distribution defined by MNase (tags plotted relative to consensus).

(C) Regions of histone crosslinking projected onto the crystal structure of one half of the nucleosomal core particle (Luger et al., 1997). Colored segments of the circle denote regions of crosslinking, centered at the indicated distance from the nucleosome dyad. Arrowhead denotes where the H3 tail emerges from between the DNA gyres.

(D) Relationship between NRL and H3 occupancy levels between the +1 and +2 nucleosomes (Table S2). Data were plotted as a 50-nucleosome moving average for all NRL > 145 bp (minimum size of a nucleosome). Data were not background subtracted.

See also Figure S1.

strikingly detailed and unexpected view of subnucleosomal structure in vivo and its relationship with the direction of transcription.

RESULTS

H3 Tails Engage Linkers

Figure 1A displays the distribution of ChIP-exo crosslinking points (exonuclease stop sites) for each of the four core histones around the 5' end of genes. None of the histones substantially crosslinked in NFRs, as expected of their general nucleosome-free status, although SAGA-regulated (Taf1-depleted) genes tended to have higher histone occupancy in promoter regions (data not shown). H2B and H4 each displayed two regions of crosslinking for each nucleosome (two vertical stripes for each

"Nuc" stripe in Figure 1A and averaged in Figure 1B). Their crosslinks corresponded to the genomic locations expected from the crystal structure of the nucleosome core particle that is centered on the genomic coordinates of nucleosome midpoints (dyads), as defined by MNase digestion (Figure 1C). H4 and H2B were about 26 and 38 bp from the dyad, respectively. H2A crosslinked broadly near linker DNA, which is consistent with in vitro studies (Shukla et al., 2011; Usachenko et al., 1994). This broad crosslinking represents two adjacent H2A from two adjacent nucleosomes. Sorting arrays by the length of the linker DNA between the +1 and +2 nucleosomes resolved H2A into separate peaks (Figure S1 available online). For brevity, we focused subsequent analyses on the +1 and +2 nucleosome positions, although equivalent conclusions can be drawn at other resolvable genic positions.

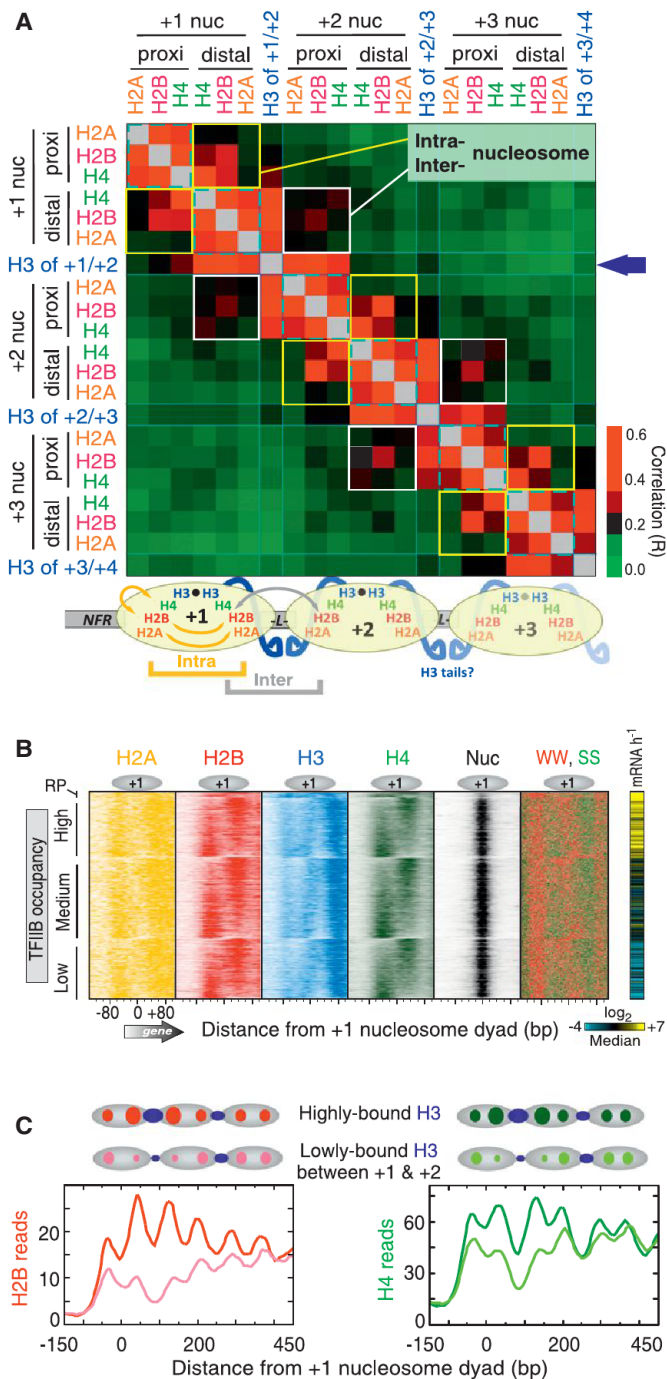


Figure 2. Asymmetry in Histone Occupancy

(A) Heatmap of correlation coefficients (R) for all pairwise combinations of NFR-proximal and NFR-distal histone occupancies at +1, +2, and +3 nucleosomes (Table S2). To simplify the patterning, correlations above 0.4 were coded with the same red color. The types of intra- and internucleosomal correlative interactions that the data suggest are illustrated below the heatmap.

(B) Histone distribution relative to the +1 nucleosome of mRNA genes (4,738 rows) orientated by TSS and sorted by H4 NFR-distal versus NFR-proximal ratios (Table S4). Rows were grouped by ribosomal protein (RP, $n = 128$) genes, then by high ($n = 1,381$), medium ($n = 1,852$), and low ($n = 1,377$) TFIIIB

Surprisingly, H3 crosslinking peaked in linker regions rather than at nucleosome dyads where the bulk of H3 resides (Figures 1A and 1B). A minor H3 peak was detected at the dyad. In the nucleosome crystal structure, the base of the H3 N-terminal tail emerges from the nucleosome core near the linker DNA (Figure 1C). To experimentally test whether H3 tails were responsible for linker crosslinking, ChIP-exo was performed in a strain where most, but not all, of the H3 tail was deleted ($\Delta 1-28$) (Morgan et al., 1991). H3-linker crosslinking was diminished in this strain (Figures 1A and 1B), suggesting that H3 tails are in close proximity to linker DNA within genic nucleosome arrays. This conclusion is further supported by *in vitro* reconstitution experiments (Zheng et al., 2005). The overall H3 pattern was not grossly altered in the H3 $\Delta 1-28$ strain (Figures 1A and 1B), indicating that amino acids 1–28 of H3 tails lack a predominant or nonredundant role in organizing nucleosomes within arrays. The remaining eight amino acids of the tail (residues 29–36) could nevertheless be involved.

Remarkably, H3 crosslinks were largely absent from the edges of NFRs adjacent to +1 nucleosomes (which might be thought of as very long linkers; Figures 1A and 1B). When arrays were sorted by +1/+2 linker length, H3 tail-linker crosslinking diminished at linkers $> \sim 30$ bp in length (i.e., where the nucleosome repeat length or NRL exceeds ~ 180 bp in Figures 1D and S1). Genes at both extremes of this linker-length distribution tended to have more dynamic nucleosomes and be of the Taf1-depleted (and thus SAGA/TATA/stress-regulated) class (right-most panels in Figure S1), which is consistent with the greater plasticity and inducibility of these genes (Huisinga and Pugh, 2004; Tirosh and Barkai, 2008). Thus, nucleosome dynamics, +1/+2 linker length, diminished H3 interactions at the +1/+2 linker, and gene inducibility appear to be linked. This may reflect an ability to mobilize (e.g., reposition or dissociate) a noncanonically positioned +1 nucleosome, which would occur adjacent to long linkers. Indeed, chromatin remodelers may access nucleosome via adjacent long linkers (Ranjan et al., 2013).

In principle, because transcription and genic arrays have directionality, H3 tail-linker interactions might arise predominantly from H3 in the adjacent upstream nucleosome or H3 in the adjacent downstream nucleosome, resulting in asymmetric substructures. Alternatively, H3 in both flanking nucleosomes might contribute similarly. We reasoned that H3 tail-linker occupancy levels should correlate most with the occupancy level of the nearest neighboring nucleosome or H4 subunit, as H3 normally interacts with H4 within a nucleosome. H3-linker occupancy correlated similarly with both flanking H4 occupancy levels (blue arrow in Figure 2A,

occupancy in the promoter region (Rhee and Pugh, 2012b). Also shown is the distribution of SS (green) and WW (red) dinucleotides (IUPAC: W = A/T, S = C/G; 4 bp bin). The right panel shows transcription frequency (Holstege et al., 1998), where yellow, blue, and black represent high, medium, and low rates, respectively.

(C) Composite distribution of histones separated out by the most highly versus the most lowly occupied H3 in the +1/+2 linker (respectively shown as dark and light colored traces, using 30th percentile cutoffs; $n = 1,455$ for each, derived from Table S3). Gray ovals demarcate nucleosome intervals. See also Figure S2.

$R_{\text{proximal}} = 0.39$ versus $R_{\text{distal}} = 0.31$) and also total flanking nucleosomal H4 levels (i.e., both H4 copies within a flanking nucleosome; Figures S2A–S2C). Similar trends were evident with H3 ($\Delta 1$ – $\Delta 8$) (Figure S2D, $R_{\text{proximal}} = 0.37$ versus $R_{\text{distal}} = 0.30$). We conclude that H3-linker interactions are largely derived from both flanking nucleosomes, although the correlations suggest there may be slightly more interactions arising from the more upstream nucleosome.

Evidence for Subnucleosomal Structures

Because the histone octamer is a fundamental unit of chromatin, we expected histone occupancies on the NFR-proximal half of a nucleosome to correlate with histone occupancies on its NFR-distal half. Indeed, occupancies did correlate between the two halves (yellow boxed areas in Figure 2A, $R_{\text{ave}} = 0.22$). However, the correlations were surprisingly modest when compared to stronger cross-correlations of the core histones located on the same half of each nucleosome (cyan dashed areas, $R_{\text{ave}} = 0.45$). Nonadjacent combinations were uncorrelated ($R_{\text{ave}} = 0.08$). Although the H3 that was crosslinking in the linker was not included in this assessment, it too behaved similarly. These relationships were also evident in plots of individual arrays (Figure 1A) and in composite plots (see Figure 2C). In general, H4–H4 cross-nucleosomal correlations were higher than those for H2A–H2A and H2B–H2B, which may reflect a tighter cross-nucleosomal linkage between H4. We interpret these correlations to reflect differential histone occupancy (or DNA crosslinking) on one half of a canonical genic nucleosome compared to the other, subject to the controls and caveats described below.

Figure 2B displays the distinct histone occupancy levels on the two halves of canonical +1 nucleosomes, where all data sets were grouped by promoter activity (TFIIB occupancy) of the associated gene, then sorted based on H4 distal/proximal ratios. Thus, where H4 occupancy was higher on the NFR-distal half of +1, the occupancy of the other histones was also higher on that half. A reciprocal relationship existed on the NFR-proximal half. These findings suggest that the histones on one half of a genic nucleosome may be more coordinated in their DNA occupancy than they are across the two halves of the same nucleosome.

About 50% of all analyzed nucleosomes showed a >2-fold differential of H2B occupancy on one half versus the other, whereas only about 12% showed the same differential with H4. These represent arbitrary thresholds, as there is a continuum of differential occupancy but subject to nonbiological biases in detection, which we address below. Thus, differential occupancy is detected between two halves of a nucleosome in a substantial number of cases. The more pronounced differential seen with H2B suggests that additional differential occupancy may exist between H3/H4 and H2A/H2B on the same half of a nucleosome. The low cross-nucleosome correlations relative to the same-side correlations at nucleosome positions +1, +2, and +3 (Figure 2A) indicate that these relationships exist broadly across multiple genic nucleosome positions.

SS richness (where S denotes G or C) in DNA promotes nucleosome assembly (Kaplan et al., 2009; Mavrich et al., 2008; Tillo and Hughes, 2009). We therefore examined the distribution of SS dinucleotides at +1 nucleosomes and found them to partially

reflect differential proximal versus distal histone occupancy (Figure 2B). This suggests that an imbalance of SS (or GC content) on one half of a nucleosome or the other may contribute to differential histone occupancy, much as it contributes to overall nucleosome occupancy. Although this also raises the question as to whether any sequence specificity in formaldehyde crosslinking or sample processing is responsible for the histone ChIP-exo patterning, our analyses of this issue in Figures S3A and S3B suggest otherwise.

In order to seek out independent evidence for differential distal versus proximal histone occupancy, we used an assay that did not involve formaldehyde or ChIP-exo. Nucleosome organization has been mapped genome-wide at high precision using an engineered cysteine at residue 47 in H4 to catalyze hydroxyl-radical cleavages in DNA near nucleosome dyads (Brogaard et al., 2012). This separates proximal from distal nucleosomal DNA. As shown in Figure 3A, H4 distal/proximal occupancy ratios for the +1 nucleosome, as determined by ChIP-exo, positively correlated with levels of distal H4S47C-cleaved fragments and negatively correlated with proximal fragments. As described in the Extended Experimental Procedures, these correlations and the abrupt changes in the trends at the extremes of the ratios are predicted outcomes of differential proximal versus distal histone occupancy.

We next used MNase ChIP-seq as a third independent probe of subnucleosomal structure. MNase probes nucleosomal particles in their native state. MNase at low activity preferentially cleaves histone-free DNA, rather than DNA within the nucleosome core. Because we were probing for subnucleosomal particles, we size-selected for library inserts in the 35–100 bp range, rather than the normal range of 120–180 bp for full nucleosomes. In this population, we detected enhanced cleavages (5' ends) starting where H3/H4 interfaces with H2A/H2B and extending through where the canonical dyad resides (Figure 3B, where the bottom trace plateaus from –30 to +30 bp from the dyad). We interpret this cleavage to reflect enhanced DNA accessibility internal to and on one side of what otherwise would be a full nucleosome. Enhanced cleavage at the expected nucleosome edge (~75 bp from the dyad) was also detected with these small fragments. Both sets of cleavages are consistent with subnucleosomal structures consisting of hexasomes (nucleosomes that lack one H2A/H2B dimer) and half-nucleosomes.

For comparison, Figure 3B also plots the distribution of subnucleosomal fragments from standard high MNase digestion (blue middle trace). This level of MNase is expected to partially nibble in from the edges of full nucleosomes up to the junction between H2A/H2B and H3/H4, due to the loose association of DNA in that region. Because full nucleosomes are quite abundant, these cleavages are expected to dominate the distribution of small fragments, thereby obscuring the presence of subnucleosomal fragments. As a control, large DNA fragments (120–180 bp) have their predominant cleavages occurring at the expected edges of full nucleosomes. Together these results are consistent with MNase detecting hexasomal and half-nucleosomal structures.

Next, we used a fourth assay, the paired-end MNase-seq data of Henikoff et al. (2011), to confirm our MNase experiments. This assay did not involve formaldehyde or ChIP. As shown in

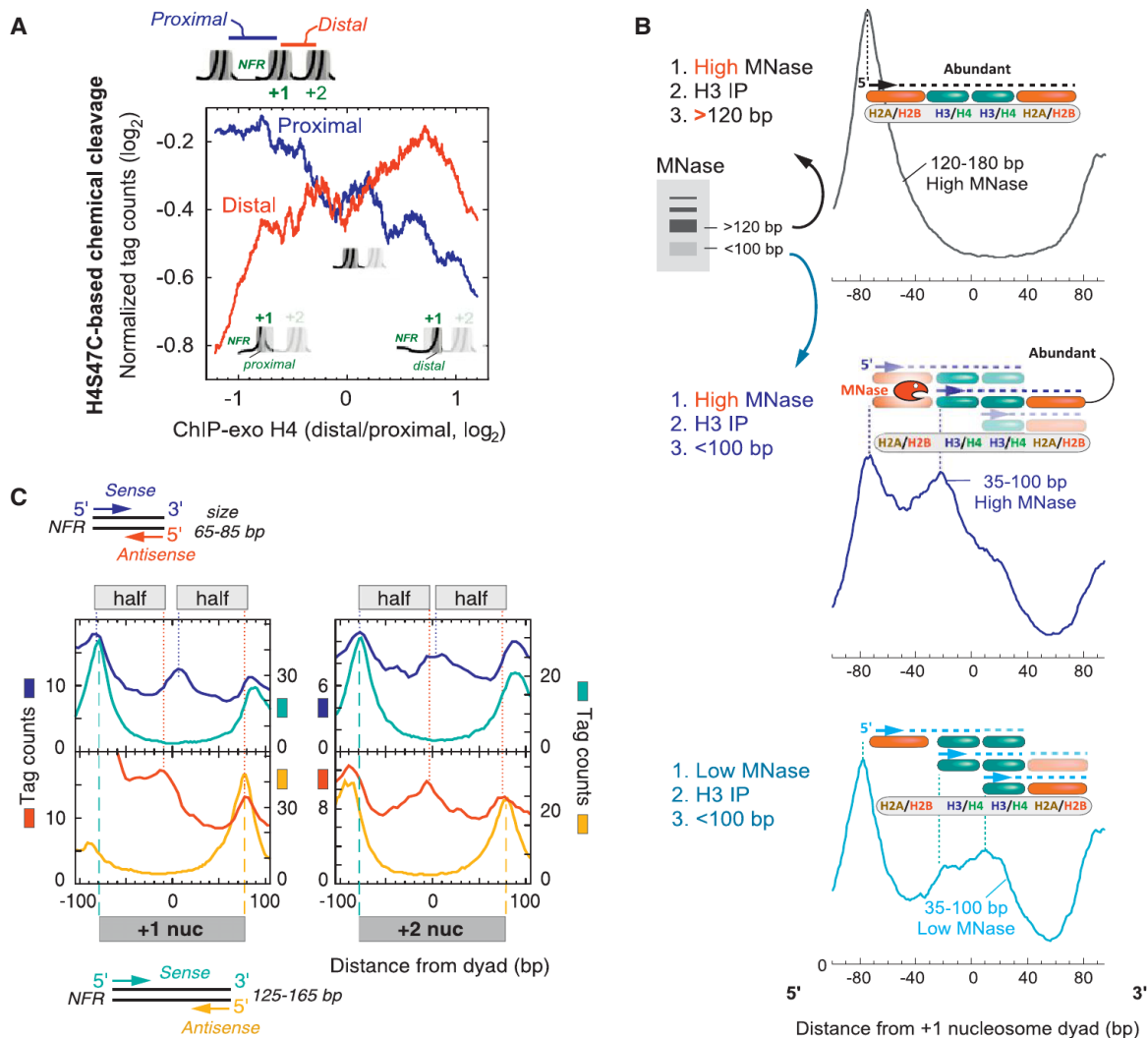


Figure 3. Evidence for Differential Proximal versus Distal Histone Occupancy using Chemical Cleavage and MNase

(A) Evidence for differential distal versus proximal histone occupancy in H4S47C-mediated chemical cleavage data (Brogaard et al., 2012). Average levels of DNA fragments (tag counts) released on the NFR-proximal (blue) or NFR-distal (red) side of 4,738 genic +1 nucleosomes are plotted as a function of H4 distal/proximal occupancy ratio at +1 nucleosomes. Data were smoothed using a 500 value moving average. See Table S5 for data processing.

(B) Cleavage at canonical nucleosomal dyads using low MNase activity. Chromatin was treated with either high or low MNase activity, H3-immunoprecipitated, then size-selected in the indicated range. The average distribution of unshifted tag 5' ends around +1 nucleosomes ($n = 4,738$) was orientated from left to right in the 5' to 3' direction (regardless of strand). Note that plotting strands separately (sense versus antisense with respect to the direction of transcription) produces essentially identical plots within the relevant +1 region, when plotted in the 5' to 3' direction (not shown). Tag counts are normalized, and thus their vertical scales are not comparable between traces. Above each trace are illustrated interpretations of the peaks. The interpretations were constrained by the experimental design so that properly sized fragments spanned H3 and included at least two dimer sets of histones. See Table S5 for data processing.

(C) Composite distribution of MNase cleavage sites (paired-end tag 5' ends) reported by Henikoff et al. (2011), plotted relative to the +1 and +2 nucleosome midpoints (Table S2) of annotated mRNA genes ($n = 4,738$) with respect to TSS orientation. Full nucleosome (125–165 bp, cyan/orange) and subnucleosome (65–85 bp, blue/red) were computationally size-selected, and their occupancy (tag counts) plotted. Cyan/blue vertical lines indicate peak 5' ends on the sense strand, and red/orange vertical lines indicate peak 5' ends on the antisense strand of paired-end reads. Data show DNA solubilized with low MNase activity (2.5 min digestion).

See also Figure S3.

Figure 3C, cleavages were again detected in the dyad region among the population of small DNA fragments. Based on the mechanism by which MNase cleaves DNA, we suggest that in these instances the DNA is lifted off of one half or quarter of the histone core where it is accessible to MNase or alternatively

results from differential distal versus proximal histone occupancy. This implies that at least some nucleosomes at the 5' end of genes (other regions not excluded) have at least one half (demarcated by the dyad) that is intact and the other half or quarter that is disassembled.

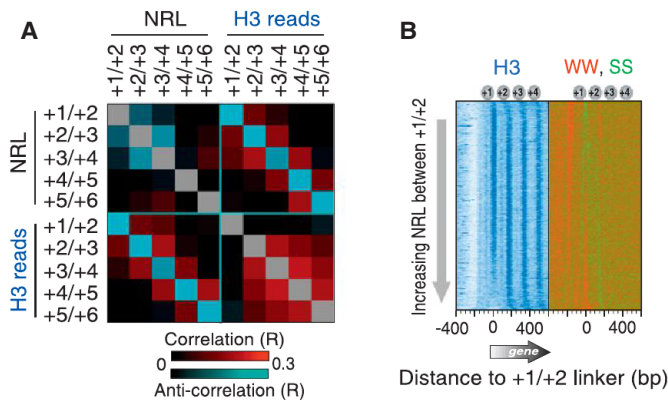


Figure 4. Sequence-Based Localized Nucleosome Positioning

(A) Heatmap representing all pairwise correlations (R) between NRLs and H3-linker occupancy levels at nucleosomes +1 through +6. Cyan and red represent negative and positive correlations, respectively. Calculations were based on nucleosomes at 3,194 genes longer than 1 kb (Table S2). NRL reports dyad-to-dyad distances in MNase-based maps (Zhang et al., 2011b). Similar observations were made with positions determined by ChIP-exo (not shown). (B) Occupancy levels of H3 ChIP-exo and SS/WW dinucleotide frequencies, relative to the midpoint between the +1 and +2 nucleosomes (Table S6), sorted by NRL between the +1 and +2 nucleosomes (Table S2).

The concept of alternative nucleosomal substructures was first described by Weintraub et al. in 1975: “A basic unit of chromosome structure is a tetramer containing all 4 histones” (Weintraub et al., 1975, 1976), where “half-nucleosomes” were reconstituted *in vitro* with pure histones and DNA. “Unfolded” nucleosomes have been isolated from cells, whereby the normally inaccessible H3-H3 dyad interface was found to be accessible to external mercury probes (Chen et al., 1991). Other subnucleosomal structures, including hexasomes, have been suggested (Annunziato, 2005; Black and Cleveland, 2011; Zlatanova et al., 2009). More recently, hemisomes have been suggested for centromeric nucleosomes (Dalal et al., 2007; Krassovsky et al., 2012) and have been reconstituted *in vitro* using centromeric H3 (Cse4) and canonical H3 (Furuyama et al., 2013). Our ChIP-exo data sets provide little insight into centromeric nucleosome structures, in that the four core histones were not detected at centromeric locations perhaps due to inefficient crosslinking or extraction (Figure S3C). However, adjacent, noncentromeric subnucleosomal structures were detected. Partial nucleosomes appear enriched at two positions to the right of centromeres, although full nucleosomes also were nearby. To the left of the centromeres, full nucleosomes predominated.

Histone Coordination between Nucleosomes

Less H3 crosslinking within the +1/+2 linker was accompanied by a parallel decrease in flanking histone occupancy (i.e., the NFR-distal half of +1 and NFR-proximal half of +2; Figure 2C, also evident in Figure 1A). Histone occupancies on the two flanks of a linker were as correlated on average as they were across two halves of the same nucleosome (Figure 2A, white versus yellow boxed areas, $R_{\text{ave.}} = 0.20$ versus 0.22, respectively). These averaged values were essentially the same at all nucleosome posi-

tions, where examined (+1, +2, +3). Thus, there appears to be coordination in histone occupancy between two halves (or parts thereof) of adjacent nucleosomes toward a shared linker. In a slight contrast, H4 was more correlated within a nucleosome than between nucleosomes. Speculatively, this might be due to the presence of H3/H4 tetramers in addition to other substructures.

NRLs (distances between MNase-defined nucleosome midpoints) anticorrelated with flanking NRLs and also anticorrelated with underlying H3 tail-linker occupancy levels (Figures 4A and S1). Thus, despite generally uniform positioning imposed by chromatin remodelers (Zhang et al., 2011b), certain nucleosome neighbors gravitate toward each other (illustrated in Figure 7, bottom panel). Longer linkers (NRLs) arise on their other flank, and this is coupled to histone depletion of the nucleosomal halves that abut these long linkers. This is consistent with longer linkers having higher histone exchange rates (Figures 1A and S1) and diminished H3-tail interactions. Given that linker length is measured using full nucleosomes (regardless of histone occupancy), their measurement should not be influenced by histone depletion. Remarkably, WW enrichment (where W denotes A or T) “painted” the linker-length landscape (Figure 4B, WW, SS panel). Thus, linker WW nucleotides may promote deviations from uniform positioning established by remodelers and in doing so enhance occupancy dynamics of adjacent histones.

H3K36me3 Negatively Regulates Linker Interactions

H3 tails are methylated (me) and acetylated (ac) in a genome-wide location-specific manner so as to potentially regulate nucleosomal arrays and transcription. We examined whether H3 modifications alter H3-linker interactions by conducting ChIP-exo after immunoprecipitation with histone modification-specific antibodies. H3K4me3 and H3K79me2/3 were enriched at their previously published array positions (i.e., nucleosome positions +1, +2, +3 for H3K4me3 and at all positions for H3K79me2/3) in a transcription-linked manner (Figures 5A–5D). In addition, linker crosslinking patterns were consistent with expectations from MNase-based maps. Therefore, these marks had no overt effect on H3-linker interactions.

In contrast, H3K36me3 had a markedly less coherent ChIP-exo pattern (Figures 5A, 5B, and S4A), despite recapitulating the known array asymmetry (i.e., depletion at the 5' end of genes). This was surprising because MNase-based maps of nucleosome cores having H3K36me3 display very robust array patterning (Figure 5C) (Zhang et al., 2011c). K36 is located at the base of the H3 tail where it emerges between the DNA gyres of the nucleosome core (Luger et al., 1997). The diffuse pattern of crosslinking associated with H3K36me3 suggests that although H3K36me3 nucleosome cores are well-positioned in genic arrays (based on MNase maps), this mark is inhibitory to H3 tail-linker crosslinking. Conceivably, K36me3 might alter the trajectory of the H3 tail as it emerges from the DNA gyres, as reported for tail mutants (Ferreira et al., 2007), or bind histone-modifying/remodeling enzymes such as the Rpd3S histone deacetylase complex or the ISW1 complex (Carrozza et al., 2005; Keogh et al., 2005) with the result of blocking H3 tail-linker interactions. A more trivial explanation may be that K36me3 renders K36 less reactive to formaldehyde, if indeed K36 is the major

point of crosslinking. To experimentally test this, we performed ChIP-exo on H3 containing alanine instead of lysine at position 36 (K36A mutant). This mutant H3 appeared to crosslink normally to linker DNA (Figure S4B), indicating that K36 is not the predominant H3 crosslink to linker DNA, although its methylation alters the potential of the tail to crosslink to linkers.

We examined linker histone H1 and found its crosslinking pattern to be almost identical to that of H3 (Figures 5A, 5B, and 5E). To test whether H1 regulates H3-linker interactions, we deleted H1 (*hho1Δ* strain). However, we observed no effect on H3-linker interactions (Figure 5F) nor any effects on nucleosome organization. Thus H1 does not play a widespread or nonredundant role in organizing nucleosomal arrays in *Saccharomyces*.

Asymmetry of H3K9ac, H2Bub, and H2A.Z

Nucleosomal arrays that encompass genes have asymmetry, as a whole, with respect to transcription-linked histone modifications, being distinct at the 5' ends of genes compared to internal and 3' locations (Henikoff, 2008; Rando and Ahmad, 2007). This asymmetry is an integral part of the transcription cycle. Because RNA polymerase II makes distinct approaches to the proximal versus distal sides of these nucleosomes, we examined whether this might be reflected in asymmetric deposition of histone marks and variants. Remarkably, at highly but not lowly transcribed genes, we found H3K9ac to be enriched primarily on the NFR-proximal half of the +1 nucleosome (Figures 5B and 5G). This is the half of the +1 nucleosome where H3 crosslinking was almost nonexistent in the general population. Indeed, a similar transcription-linked enrichment at the NFR-proximal side of +1 was not observed for H3 or for other transcription-linked H3 marks. This is consistent with the notion that the K9ac mark is transient, and that most H3 is unacetylated at H3K9 even at highly transcribed genes. Thus, H3K9ac rather than transcription per se may be more directly linked to H3-tail contacts with the edge of NFRs.

H2BK123 is located within the core C-terminal α helix of H2B, and its ubiquitylation (ub) is linked to the transcription cycle (Batta et al., 2011; Fleming et al., 2008; Pavri et al., 2006). Like H3K9ac, H2Bub was highly enriched on the NFR-proximal half of the +1 nucleosome in a transcription-linked manner (Figures 5B and 5H). In light of this and the prior observation that loss of ubiquitylation results in accumulation of RNA polymerase II at promoters (Batta et al., 2011), we speculate that H2B ubiquitylation at the NFR-proximal half of the +1 nucleosome facilitates the movement of RNA polymerase II into gene bodies. Further into gene bodies, crosslinking of additional H2Bub occurred in linkers and was contributed by both flanking H2B (Figure S4C), as seen with H3. Its basis is currently unclear, although preferential crosslinking of ubiquitylated H2B to H3 could generate such a pattern.

The apparent nucleosome asymmetry of histone marks could be an indirect consequence of differential histone occupancy (Figure 6A). To address this possibility, we normalized the level of each mark to the underlying occupancy level of the relevant histone (e.g., H3K9ac/H3). We then calculated the log₂ occupancy ratio at the distal half of the +1 nucleosome to its proximal half and plotted it as a function of H4 distal/proximal ratio. Accordingly, H3K9ac retained its overall preference for the

NFR-proximal half of +1, as indicated by the black trace in Figure 6B being below zero and having zero slope. One caveat is that H3 and H3K9ac on the distal half were not positionally resolved from the proximal half of the +2 nucleosome.

When normalized to H2B, ubiquitylated H2B remained biased toward the NFR-proximal half of the +1 nucleosome (red trace being generally below zero in Figure 6B). In addition, it displayed a tendency toward occupying the half that was most depleted of histones (negative slope of the trace). These effects were accentuated at highly transcribed genes (not shown). As depletion likely reflects dynamics, we infer that higher ubiquitylation density reflects nascent histone assembly occurring on the depleted side, which is what H2B ubiquitylation is thought to promote.

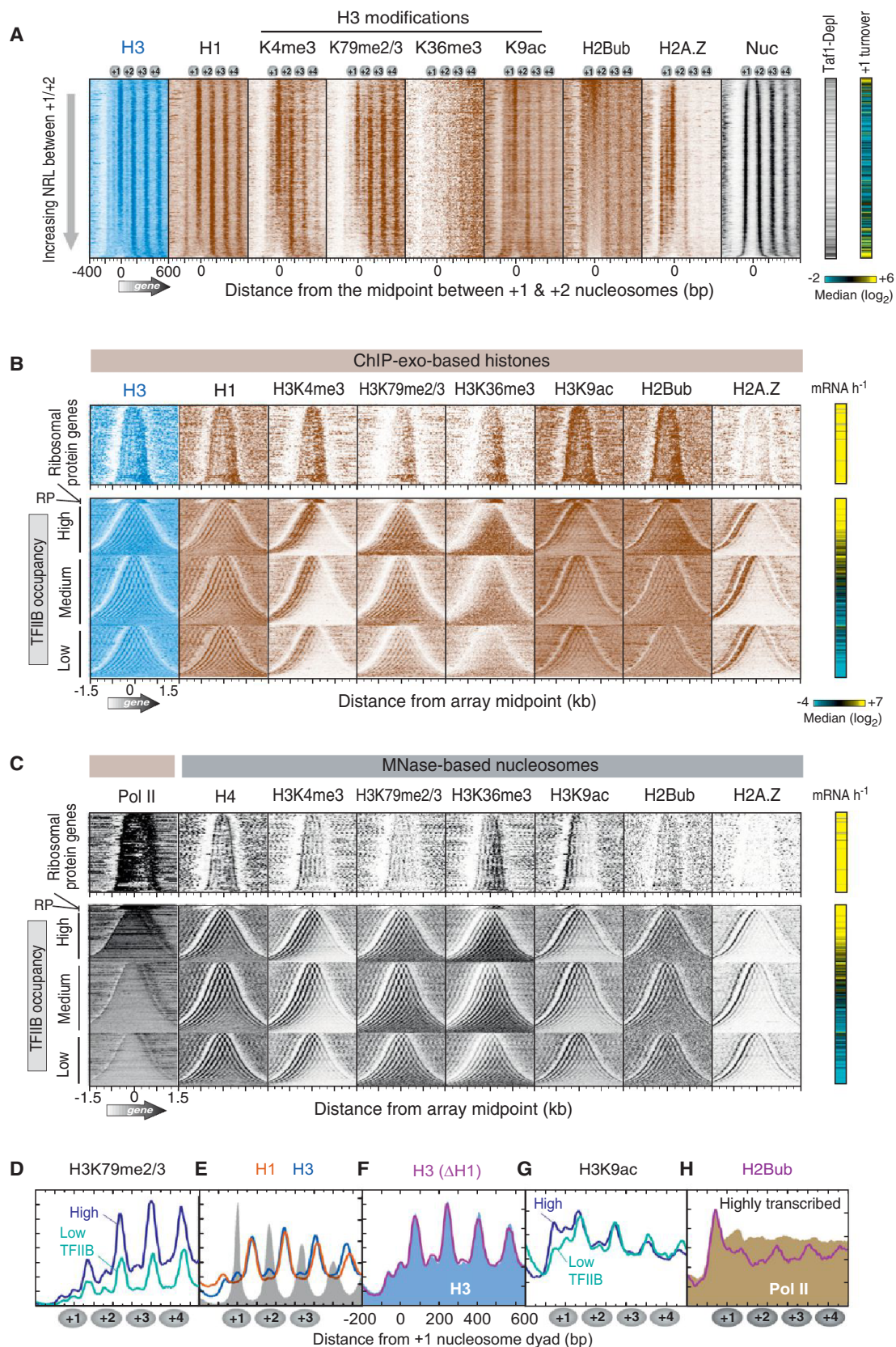
As expected, histone variant H2A.Z was highly enriched at the +1 nucleosome position (Figures 5A and 5B) (Albert et al., 2007) and was asymmetrically placed toward the NFR-distal half of the +1 nucleosome (Figures 5A, 6C, and 6D). Compared to H2A, H2A.Z was preferentially enriched where H4 was enriched (positive H2A.Z/H2A slope in Figure 6B), thereby linking the presence of H2A.Z to stably occupied histones. An exception was at long +1/+2 linkers, where the trend was reversed. This is consistent with the SWR1/SWR-C complex requiring long DNA to deposit H2A.Z into nucleosomes (Ranjan et al., 2013; Yen et al., 2013). We conclude that the transcription machinery generally encounters H2A.Z on the distal half of the +1 nucleosome. Such heterotypic nucleosomes are intrinsically unstable (Bönisch and Hake, 2012) and thus might facilitate the passage of RNA polymerase II (Weber et al., 2014).

DISCUSSION

Subnucleosomal Structures Suggest Mechanisms for Nucleosome Dynamics

The results presented here provide insight into potential mechanisms of nucleosome instability at the 5' ends of genes that differs from the canonical view. Collectively, the data show that histone occupancies on one half of a nucleosome dyad are more strongly coordinated with each other than histone occupancy across the two sides of the dyad (Figure 7, bottom panel). The effect is stronger for H2A/H2B than for H3/H4, leading us to surmise that both hexasomes (two copies of H3/H4 and one copy of H2A/H2B) and half-nucleosomes (one copy of H2A/H2B/H3/H4) exist in vivo, in addition to the more abundant standard nucleosomes. This observation is consistent with early views of in vivo nucleosomes (Weintraub et al., 1975, 1976) and with biochemical studies that reconstitute such half-nucleosomes on DNA (Furuyama et al., 2013). Such partial nucleosomes may not be present or evident in typical in vitro reconstitution studies, possibly due to missing factors (e.g., chaperones and/or DNA sequence) and/or the use of assays that are unsuitable for their detection in a subpopulation.

In support of the physiological importance of subnucleosomal structures, nucleosome positions that were associated with noncanonical linker lengths, which are linked to differential distal versus proximal histone occupancies, had distinct properties. They tended to have higher histone turnover and were associated with Taf1-depleted/SAGA-regulated genes. Differential histone occupancy was also associated with distinctive densities of



(legend on next page)

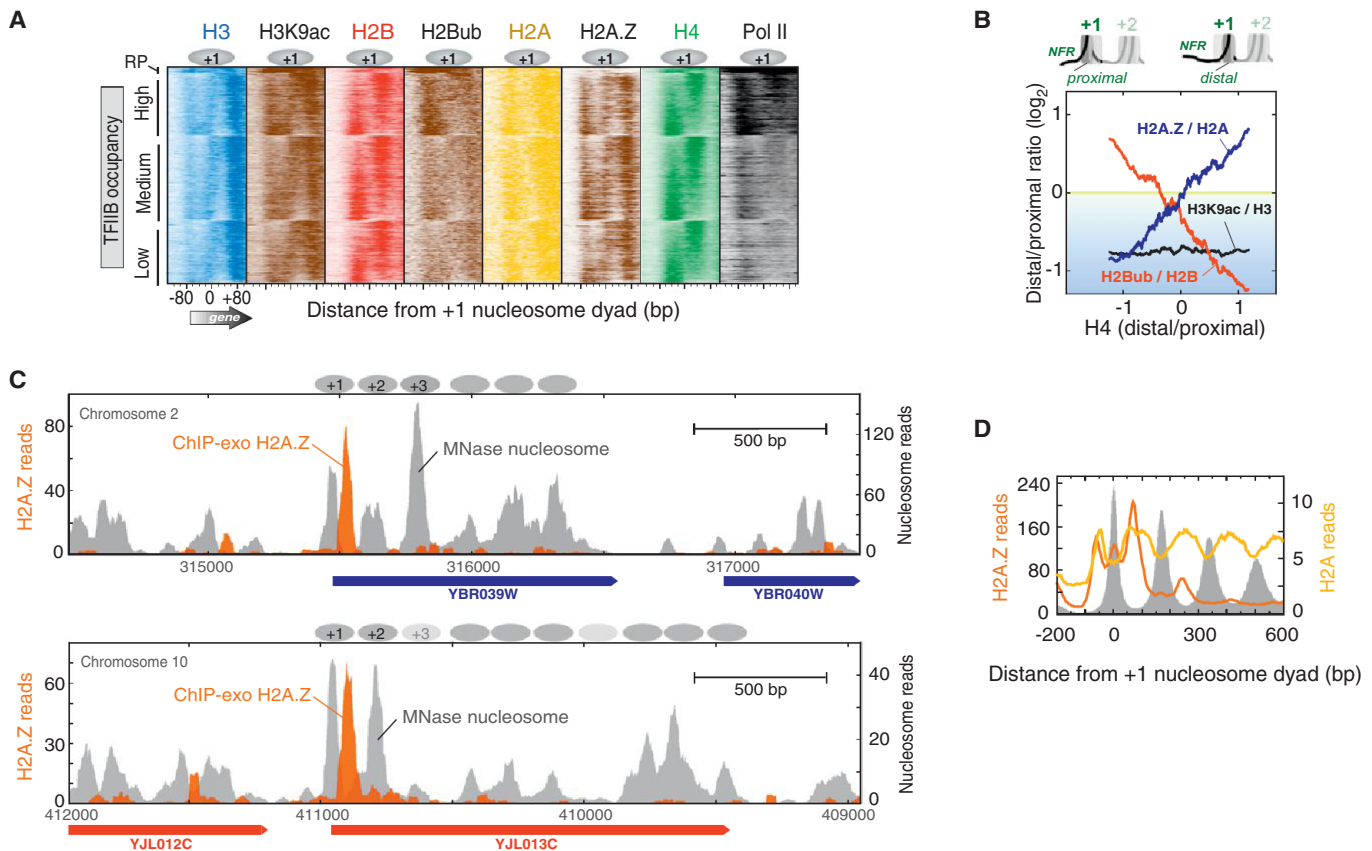


Figure 6. Asymmetric Density of Histone Marks and Variant H2A.Z

(A) Proximal versus distal occupancy of histone marks. Data were plotted as described in Figure 2B.

(B) Shown is the distal/proximal occupancy ratio of the indicated histone marks or variant at +1 nucleosomes ($n = 4,738$), after normalizing to (dividing by) the relevant underlying core histone occupancy (intervals designated in Table S2). Ratios were \log_2 transformed, sorted by H4 distal/proximal ratio (x axis), and smoothed using a 500-nucleosome moving average.

(C) Example of asymmetrically placed H2A.Z at the NFR-distal half of the +1 nucleosome at specific loci. Orange filled plot shows the distribution of H2A.Z-crosslinking sites (raw sequencing tags) measured by ChIP-exo. Gray filled plot shows nucleosome midpoints detected by MNase ChIP-seq.

(D) Composite distribution H2A.Z (orange trace) and H2A (yellow trace) detected by ChIP-exo, relative to the +1 nucleosome midpoint at all mRNA genes. Gray fill indicates nucleosome midpoint distribution detected by MNase ChIP-seq.

histone marks. Thus, regardless of its structural basis, differential proximal versus distal histone occupancy is associated with distinct functional properties compared to all other nucleosomes at the same relative position. Differential occupancy was not correlated with transcription frequency of the underlying mRNA

gene, which indicates that it is not necessarily linked to transcription. Whether it is linked to noncoding transcription or other types of genomic regulation remains to be determined.

In addition to previously published works, evidence for half-nucleosomes in this study comes from several different

Figure 5. Subnucleosomal Organization of Histone Marks and Variants

(A) Occupancy levels of histones and their marks relative to the +1/+2 linker and sorted by linker length ($n = 4,738$ genes; Table S7). The right panels demarcate Taf1-depleted genes (black lines) (Rhee and Pugh, 2012b) and histone turnover rate of the +1 nucleosome region (yellow indicates “hot” nucleosomes) (Dion et al., 2007).

(B) Same as (A) except that entire genic arrays are shown (5' to 3' from left to right). Arrays are sorted by array length and grouped by ribosomal protein (RP) genes (including an expanded vertical view), and the remaining by TFIIB occupancy (Rhee and Pugh, 2012b). The right panel shows transcription frequency (Holstege et al., 1998).

(C) Similar as (B), except measured by MNase ChIP-seq from previous studies (Albert et al., 2007; Batta et al., 2011; Zhang et al., 2011a, 2011b). Pol II denotes RNA polymerase II measured by ChIP-exo.

(D–H) Composite distribution of the indicated histones or marks relative to the +1 nucleosome midpoint at genes having TFIIB promoter occupancy in the top and bottom 30% (Rhee and Pugh, 2012b). (E) and (F) are for all genes. (F) H3 in a wild-type strain is shown as a blue fill, and H3 in an *hho1Δ* strain as a magenta trace. (H) H2Bub (magenta) and Pol II (brown fill) are for genes having the top 30% of TFIIB occupancy.

See also Figure S4.

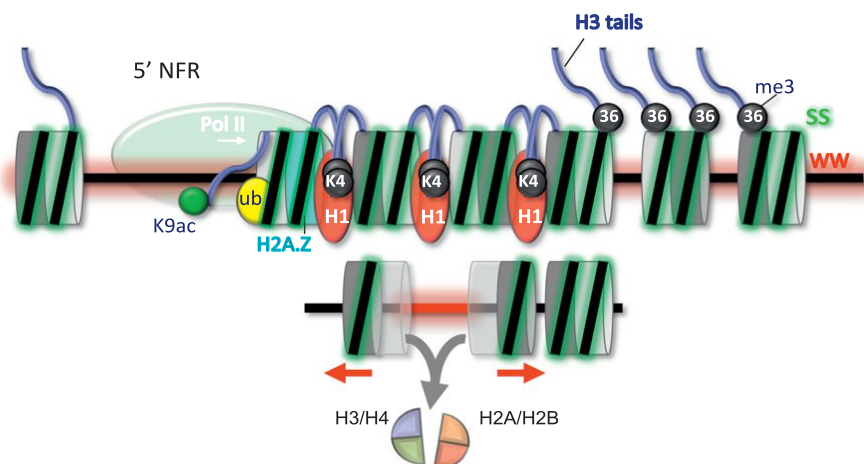


Figure 7. Composite Model of Nucleosomal Arrays at Genes

The nucleosome-free promoter region (5' NFR) is depicted with RNA polymerase II. Within arrays, the two sides of a nucleosome are depicted as two-toned gray disks, fused to form nucleosomes. In the upper panel, histone marks, variants, and WW/SS dinucleotide enrichment are depicted by different colors. "K9ac," "ub," "K4," and "36" denote H3K9ac, H2Bub, H3K4me3, and H3K36me3, respectively. These features are placed at their predominant locations within arrays. The lower panel illustrates longer WW-rich linkers (red) that are linked to histone destabilization on the linker-flanking half of each nucleosome (symbolized by transparency). The red arrows indicate that those (sub)nucleosomes are closer than normal to their adjacent nucleosomes.

assays: (1) correlated distal/proximal occupancy ratios of all four core histones, measured by the ChIP-exo assay performed on in vivo crosslinked chromatin; (2) differential occupancy measured by H4S47C-based chemical cleavage of noncrosslinked chromatin; (3) MNase cleavage at dyads of native crosslinked chromatin; (4) MNase cleavage at dyads using noncrosslinked chromatin and paired-end sequencing from a different lab; and (5) an enrichment of DNA sequences (GC) that favor nucleosome formation on the side where histone occupancy is the highest. A number of control analyses indicated that DNA sequence-based biases in either crosslinking or DNA sequencing was not a trivial explanation for differential occupancy. However, we cannot unequivocally rule this possibility out.

Our conclusions on half-nucleosomes run counter to the current view that nucleosome assembly/disassembly proceeds through an H3/H4 tetramer (dimer of dimers) intermediate. Conceivably, however, nucleosome assembly/disassembly might proceed by multiple pathways. For example, *ab initio* nucleosome reassembly in the wake of DNA or RNA polymerase might proceed via the classical pathway of chaperone-assisted assembly of H3/H4 tetramers (dimer of dimers) followed by chaperone-assisted assembly of H2A/H2B dimers. However, chromatin that is intrinsically dynamic in the absence of a passing polymerase might exchange histones via hexasome and half-nucleosome intermediates. In this way, a hexasome might lose either a single H2A/H2B dimer or a single H3/H4 dimer. Chaperones, chromatin remodelers, DNA/RNA polymerases, histone variants/modifications, and the underlying DNA sequence could influence pathway selection in response to environmental cues.

Having more dynamic nucleosomes, for example near the 5' ends of genes, might alter accessibility to transcription factor-binding sites or TSSs, and this includes creating alternative coding and noncoding TSSs. The class of genes that are particularly associated with subnucleosomal structures tend to also have abnormal +1/+2 linker lengths. These genes tend to be SAGA regulated and TATA containing and thus subject to a wide range of regulation.

Chromatin remodelers play a predominate role in organizing nucleosomes into uniform arrays. However, the work presented

here and elsewhere suggests that the underlying DNA sequence may help maintain a prescribed organizational state at certain genes that differs somewhat from the canonical pattern that remodelers offer (Kaplan et al., 2009; Mavrich et al., 2008; Tillo and Hughes, 2009). These genes tend to have dynamic chromatin, and this fluidity may allow the underlying DNA sequence to reposition certain nucleosomes away from remodeler-imposed spacing uniformity. We observed coordinated histone occupancy between two halves of adjacent nucleosomes that share a common linker. Nucleosome repositioning may alter this relationship resulting in altered nucleosome stability.

Nucleosome splitting at the dyad, as one possible interpretation presented here, might explain the confounding biophysical observation that mechanical disruption of nucleosomes with optical traps produces a large cooperative tension transition that cannot readily be explained by gradual unwrapping of DNA from the histone octamer surface (Brower-Toland et al., 2002). Instead the abrupt transition may reflect the splitting of a nucleosome into two halves.

Nucleosomes Are Asymmetric with Respect to Transcription

Our findings suggest that RNA polymerase II first encounters a unique and asymmetric +1 nucleosome (Figure 7). On the proximal face, which encounters polymerase first, these asymmetric features include H3-tail interactions with the edge of the NFR that are coupled to H3K9 acetylation. Further into gene bodies, H3 tails engage linkers, without dependence on H3K9ac, but their purpose remains unclear. Loss of most (but not all) of the H3 tail has surprisingly little impact on nucleosome organization, as does loss of histone H1. Methylation of H3 lysines 4 and 79 (with the former on the H3 tail) has little apparent impact on H3 tail-linker interactions. However, H3K36 methylation substantially delocalizes the crosslinking of H3 tails, such that it loses substantial specificity for linkers. K36 methylation might alter the trajectory of the H3 tail away from linkers, or the H3K36 modification might bind known regulatory factors such as Rpd3S or ISW1 complexes (Carrozza et al., 2005; Smolle et al., 2012), thereby preventing crosslinking.

Back at the +1 nucleosome, selective placement of H2B ubiquitylation on the NFR-proximal face may be important for allowing RNA polymerase II passage as well as promoting nucleosome reassembly in the wake of transcription. The presence of H2A.Z on the distal face of the +1 nucleosome may further facilitate the passage of polymerase, through destabilization of the promoter-proximal half of the nucleosome. This subnucleosomal asymmetry as well as overall array asymmetry may be applicable to multicellular eukaryotes including humans because histones, nucleosome organization, and modifications are highly conserved across species.

EXPERIMENTAL PROCEDURES

ChIP-Exo Assay

Saccharomyces strains (listed in Table S1) were grown in rich media and subjected to formaldehyde crosslinking, then processed through the ChIP-exo assay (Rhee and Pugh, 2012a), using either SOLiD or Illumina adaptors. Briefly, cells were disrupted, and chromatin pellets were isolated and then solubilized and fragmented by sonication. Fragmented chromatin was then subject to immunoprecipitation using magnetic bead-conjugated antibodies directed either against TAP-tagged histones or directly against histones or their modifications. After washing the beads to remove unbound proteins and DNA, and while still on the beads, the immunoprecipitates were polished, A-tailed, and ligated to an appropriate sequencing library adaptor. Samples were then subjected to lambda exonuclease digestion, which processively removes nucleotides from 5' ends of double-stranded DNA until blocked by a protein-DNA crosslink induced by formaldehyde treatment. The result is single-stranded DNA, which was then eluted from the magnetic beads and converted to double-stranded DNA by primer annealing and extension. A second sequencing adaptor was then ligated to exonuclease treated ends, then PCR amplified, gel purified, and sequenced.

MNase-Seq

Figure 3B experiments involved MNase-seq, where crosslinked chromatin was treated with 240 units of MNase for either 5 or 25 min (low versus high), then subjected to H3 immunoprecipitation, Illumina library construction, gel size selection, and deep sequencing. Libraries were sequenced by an Applied Biosystems 5500xl SOLiD System, Illumina HiSeq2000, or Illumina NextSeq500. Sequencing tags were mapped to the reference yeast genome obtained from *Saccharomyces* Genome Database (R55-10-Nov-2006).

Data Analysis

Occupancy levels (tag counts) for various histone positions were typically plotted relative to genomic reference points of MNase-derived nucleosome dyads (listed in Table S2) and summed within the intervals specified in Table S2. Where occupancy correlations between data sets are reported, the Excel function "correl" was used. Tables S3, S4, S5, S6, and S7 present the underlying values and calculations used in the figures. Array midpoints represent the coordinate located half-way between the dyad of the +1 nucleosome and the dyad of the terminal genic nucleosome. See Extended Experimental Procedures for detailed experimental procedures, analyses methods, and rationale.

ACCESSION NUMBERS

Sequencing data are available at NCBI Sequence Read Archive under accession number SRA059355.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, four figures, and seven tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2014.10.054>.

ACKNOWLEDGMENTS

We thank Yunfei Li, Rohit Reja, and William Lai for bioinformatic support, Matthew Rossi for sharing unpublished ChIP input DNA data, and members of the Pugh laboratory, the Penn State Center for Eukaryotic Gene Regulation, and Philipp Korber for valuable discussions. We are grateful to Bongsoo Park for computational assistance. National Institutes of Health grant HG004160 supported this work. B.F.P. has a financial interest in Peconic, LLC, which utilizes the ChIP-exo technology implemented in this study and could potentially benefit from the outcomes of this research.

Received: April 4, 2014

Revised: July 19, 2014

Accepted: October 13, 2014

Published: December 4, 2014

REFERENCES

- Albert, I., Mavrich, T.N., Tomsho, L.P., Qi, J., Zanton, S.J., Schuster, S.C., and Pugh, B.F. (2007). Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature* **446**, 572–576.
- Annunziato, A.T. (2005). Split decision: what happens to nucleosomes during DNA replication? *J. Biol. Chem.* **280**, 12065–12068.
- Batta, K., Zhang, Z., Yen, K., Goffman, D.B., and Pugh, B.F. (2011). Genome-wide function of H2B ubiquitylation in promoter and genic regions. *Genes Dev.* **25**, 2254–2265.
- Black, B.E., and Cleveland, D.W. (2011). Epigenetic centromere propagation and the nature of CENP-a nucleosomes. *Cell* **144**, 471–479.
- Bönisch, C., and Hake, S.B. (2012). Histone H2A variants in nucleosomes and chromatin: more or less stable? *Nucleic Acids Res.* **40**, 10719–10741.
- Brogaard, K., Xi, L., Wang, J.P., and Widom, J. (2012). A map of nucleosome positions in yeast at base-pair resolution. *Nature* **486**, 496–501.
- Brower-Toland, B.D., Smith, C.L., Yeh, R.C., Lis, J.T., Peterson, C.L., and Wang, M.D. (2002). Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA. *Proc. Natl. Acad. Sci. USA* **99**, 1960–1965.
- Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P., and Workman, J.L. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**, 581–592.
- Chen, T.A., Smith, M.M., Le, S.Y., Sternglanz, R., and Allfrey, V.G. (1991). Nucleosome fractionation by mercury affinity chromatography. Contrasting distribution of transcriptionally active DNA sequences and acetylated histones in nucleosome fractions of wild-type yeast cells and cells expressing a histone H3 gene altered to encode a cysteine 110 residue. *J. Biol. Chem.* **266**, 6489–6498.
- Dalal, Y., Furuyama, T., Vermaak, D., and Henikoff, S. (2007). Structure, dynamics, and evolution of centromeric nucleosomes. *Proc. Natl. Acad. Sci. USA* **104**, 15974–15981.
- Dion, M.F., Kaplan, T., Kim, M., Buratowski, S., Friedman, N., and Rando, O.J. (2007). Dynamics of replication-independent histone turnover in budding yeast. *Science* **315**, 1405–1408.
- Ferreira, H., Somers, J., Webster, R., Flaus, A., and Owen-Hughes, T. (2007). Histone tails and the H3 alphaN helix regulate nucleosome mobility and stability. *Mol. Cell. Biol.* **27**, 4037–4048.
- Fleming, A.B., Kao, C.F., Hillyer, C., Pikaart, M., and Osley, M.A. (2008). H2B ubiquitylation plays a role in nucleosome dynamics during transcription elongation. *Mol. Cell* **31**, 57–66.
- Furuyama, T., Codomo, C.A., and Henikoff, S. (2013). Reconstitution of hemisomes on budding yeast centromeric DNA. *Nucleic Acids Res.* **41**, 5769–5783.
- Henikoff, S. (2008). Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat. Rev. Genet.* **9**, 15–26.

- Henikoff, J.G., Belsky, J.A., Krassovsky, K., MacAlpine, D.M., and Henikoff, S. (2011). Epigenome characterization at single base-pair resolution. *Proc. Natl. Acad. Sci. USA* **108**, 18318–18323.
- Holstege, F.C., Jennings, E.G., Wyrick, J.J., Lee, T.I., Hengartner, C.J., Green, M.R., Golub, T.R., Lander, E.S., and Young, R.A. (1998). Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**, 717–728.
- Huisinga, K.L., and Pugh, B.F. (2004). A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Mol. Cell* **13**, 573–585.
- Jiang, C., and Pugh, B.F. (2009). Nucleosome positioning and gene regulation: advances through genomics. *Nat. Rev. Genet.* **10**, 161–172.
- Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Tillo, D., Field, Y., LeProust, E.M., Hughes, T.R., Lieb, J.D., Widom, J., and Segal, E. (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**, 362–366.
- Keogh, M.C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J., et al. (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* **123**, 593–605.
- Kornberg, R.D., and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* **98**, 285–294.
- Krassovsky, K., Henikoff, J.G., and Henikoff, S. (2012). Tripartite organization of centromeric chromatin in budding yeast. *Proc. Natl. Acad. Sci. USA* **109**, 243–248.
- Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260.
- Luger, K., Dechassa, M.L., and Tremethick, D.J. (2012). New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat. Rev. Mol. Cell Biol.* **13**, 436–447.
- Mavrich, T.N., Ioshikhes, I.P., Venters, B.J., Jiang, C., Tomsho, L.P., Qi, J., Schuster, S.C., Albert, I., and Pugh, B.F. (2008). A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res.* **18**, 1073–1083.
- Morgan, B.A., Mittman, B.A., and Smith, M.M. (1991). The highly conserved N-terminal domains of histones H3 and H4 are required for normal cell cycle progression. *Mol. Cell. Biol.* **11**, 4111–4120.
- Pavri, R., Zhu, B., Li, G., Trojer, P., Mandal, S., Shilatfard, A., and Reinberg, D. (2006). Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* **125**, 703–717.
- Rando, O.J., and Ahmad, K. (2007). Rules and regulation in the primary structure of chromatin. *Curr. Opin. Cell Biol.* **19**, 250–256.
- Ranjan, A., Mizuguchi, G., FitzGerald, P.C., Wei, D., Wang, F., Huang, Y., Luk, E., Woodcock, C.L., and Wu, C. (2013). Nucleosome-free region dominates histone acetylation in targeting SWR1 to promoters for H2A.Z replacement. *Cell* **154**, 1232–1245.
- Rhee, H.S., and Pugh, B.F. (2011). Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell* **147**, 1408–1419.
- Rhee, H.S., and Pugh, B.F. (2012a). ChIP-exo method for identifying genomic location of DNA-binding proteins with near-single-nucleotide accuracy. *Curr. Protoc. Mol. Biol. Chapter 21*, Unit21.24.
- Rhee, H.S., and Pugh, B.F. (2012b). Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature* **483**, 295–301.
- Segal, E., and Widom, J. (2009). What controls nucleosome positions? *Trends Genet.* **25**, 335–343.
- Shukla, M.S., Syed, S.H., Goutte-Gattat, D., Richard, J.L., Montel, F., Hamiche, A., Travers, A., Faivre-Moskalenko, C., Bednar, J., Hayes, J.J., et al. (2011). The docking domain of histone H2A is required for H1 binding and RSC-mediated nucleosome remodeling. *Nucleic Acids Res.* **39**, 2559–2570.
- Smolle, M., Venkatesh, S., Gogol, M.M., Li, H., Zhang, Y., Florens, L., Washburn, M.P., and Workman, J.L. (2012). Chromatin remodelers Isw1 and Chd1 maintain chromatin structure during transcription by preventing histone exchange. *Nat. Struct. Mol. Biol.* **19**, 884–892.
- Tillo, D., and Hughes, T.R. (2009). G+C content dominates intrinsic nucleosome occupancy. *BMC Bioinformatics* **10**, 442.
- Tirosh, I., and Barkai, N. (2008). Two strategies for gene regulation by promoter nucleosomes. *Genome Res.* **18**, 1084–1091.
- Usachenko, S.I., Bavykin, S.G., Gavin, I.M., and Bradbury, E.M. (1994). Rearrangement of the histone H2A C-terminal domain in the nucleosome. *Proc. Natl. Acad. Sci. USA* **91**, 6845–6849.
- Weber, C.M., Ramachandran, S., and Henikoff, S. (2014). Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol. Cell* **53**, 819–830.
- Weintraub, H., Palter, K., and Van Lente, F. (1975). Histones H2a, H2b, H3, and H4 form a tetrameric complex in solutions of high salt. *Cell* **6**, 85–110.
- Weintraub, H., Worcel, A., and Alberts, B. (1976). A model for chromatin based upon two symmetrically paired half-nucleosomes. *Cell* **9**, 409–417.
- Yen, K., Vinayachandran, V., and Pugh, B.F. (2013). SWR-C and INO80 chromatin remodelers recognize nucleosome-free regions near +1 nucleosomes. *Cell* **154**, 1246–1256.
- Zhang, L., Ma, H., and Pugh, B.F. (2011a). Stable and dynamic nucleosome states during a meiotic developmental process. *Genome Res.* **21**, 875–884.
- Zhang, Z., Wippo, C.J., Wal, M., Ward, E., Korber, P., and Pugh, B.F. (2011b). A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. *Science* **332**, 977–980.
- Zhang, J., McCabe, K.A., and Bell, C.E. (2011c). Crystal structures of lambda exonuclease in complex with DNA suggest an electrostatic ratchet mechanism for processivity. *Proc. Natl. Acad. Sci. USA* **108**, 11872–11877.
- Zheng, C., Lu, X., Hansen, J.C., and Hayes, J.J. (2005). Salt-dependent intra- and internucleosomal interactions of the H3 tail domain in a model oligonucleosomal array. *J. Biol. Chem.* **280**, 33552–33557.
- Zlatanova, J., Bishop, T.C., Victor, J.M., Jackson, V., and van Holde, K. (2009). The nucleosome family: dynamic and growing. *Structure* **17**, 160–171.

Proteostatic Control of Telomerase Function through TRiC-Mediated Folding of TCAB1

Adam Freund,¹ Franklin L. Zhong,² Andrew S. Venteicher,¹ Zhaojing Meng,³ Timothy D. Veenstra,³ Judith Frydman,⁴ and Steven E. Artandi^{1,2,5,*}

¹Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA

²Cancer Biology Program, Stanford University School of Medicine, Stanford, CA 94305, USA

³Laboratory of Proteomics and Analytical Technologies, Science Applications International Corporation-Frederick, National Cancer Institute at Frederick, Frederick, MD 21702, USA

⁴Department of Biology, Stanford University, Stanford, CA 94305, USA

⁵Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305, USA

*Correspondence: sartandi@stanford.edu

<http://dx.doi.org/10.1016/j.cell.2014.10.059>

SUMMARY

Telomere maintenance by telomerase is impaired in the stem cell disease dyskeratosis congenita and during human aging. Telomerase depends upon a complex pathway for enzyme assembly, localization in Cajal bodies, and association with telomeres. Here, we identify the chaperonin CCT/TRiC as a critical regulator of telomerase trafficking using a high-content genome-wide siRNA screen in human cells for factors required for Cajal body localization. We find that TRiC is required for folding the telomerase cofactor TCAB1, which controls trafficking of telomerase and small Cajal body RNAs (scaRNAs). Depletion of TRiC causes loss of TCAB1 protein, mislocalization of telomerase and scaRNAs to nucleoli, and failure of telomere elongation. DC patient-derived mutations in TCAB1 impair folding by TRiC, disrupting telomerase function and leading to severe disease. Our findings establish a critical role for TRiC-mediated protein folding in the telomerase pathway and link proteostasis, telomere maintenance, and human disease.

INTRODUCTION

Telomeres are nucleoprotein structures that protect chromosome ends and serve as substrates for the enzyme telomerase (Palm and de Lange, 2008). Telomeres shorten progressively with cell division due to incomplete replication of the lagging DNA strand, and this shortening is offset by processive elongation of telomeres by telomerase (Pfeiffer and Lingner, 2013). Impaired maintenance of telomeres contributes to the pathogenesis of many disease states, including dyskeratosis congenita, pulmonary fibrosis, aplastic anemia, and liver cirrhosis (Armanios and Blackburn, 2012). Conversely, effective maintenance of telomeres by telomerase is thought to be important in progression of human cancers (Artandi and DePinho, 2010; Hahn et al.,

1999; Horn et al., 2013; Huang et al., 2013; Killela et al., 2013). The telomerase ribonucleoprotein requires a complex series of biochemical steps to enable enzymatic function at telomeres. Efforts to develop telomerase therapeutics useful in diverse diseases require a more complete understanding of these steps and the molecules that govern them. Unbiased approaches to identify telomerase regulators in human cells have thus far been limited by difficulty in detecting telomerase components, precluding the development of genetic screens to interrogate this pathway.

Active telomerase enzyme is comprised of a catalytic core—the telomerase RNA component, TERC, and the telomerase reverse transcriptase, TERT—in addition to several additional proteins required for proper telomerase function. The biogenesis of functional telomerase is mediated by a series of maturation, assembly, and trafficking steps that take place within the nucleus, as well as within Cajal bodies—subnuclear structures devoted to RNA modification and assembly (Egan and Collins, 2012). Human TERC shares sequence motifs in common with small Cajal body RNAs (scaRNAs), which act as guides for post-transcriptional modification of splicing RNAs within Cajal bodies (Darzacq et al., 2002; Jádý et al., 2004; Zhu et al., 2004). The processes regulating telomerase assembly overlap considerably with biogenesis pathways for scaRNAs and related small nucleolar RNAs (snoRNAs), all of which share an H/ACA sequence recognized by the dyskerin core complex—dyskerin, NHP2, and NOP10. TERC stability is dependent on the dyskerin core complex and on the assembly factors NAF1, Shq1, and pontin/reptin (Egan and Collins, 2012). Incorporation of TERT protein into an RNP containing TERC and the dyskerin core complex yields a telomerase enzyme that is stable and catalytically active (Cohen et al., 2007; Mitchell et al., 1999).

After these initial stages of assembly, telomerase localizes to Cajal bodies by association with TCAB1, which recognizes the CAB box sequence common to TERC and scaRNAs (Cristofari et al., 2007; Tycowski et al., 2009; Venteicher et al., 2009). Telomerase is specifically recruited to telomeres through an interaction between TERT and the oligonucleotide/oligosaccharide binding (OB) fold of the telomere binding protein TPP1 (Abreu et al., 2010; Nandakumar et al., 2012; Sexton et al., 2012; Zhong

et al., 2012), and this step requires TCAB1 (Stern et al., 2012; Zhong et al., 2012). TCAB1 is required for trafficking of telomerase to Cajal bodies and for telomere maintenance (Venteicher et al., 2009). This requirement is highlighted by the existence of TCAB1 mutations in patients with dyskeratosis congenita (DC), a stem cell disease caused by a failure of telomere maintenance (Zhong et al., 2011). DC is characterized by an epidermal triad (oral leukoplakia, nail dystrophy, and skin pigmentation abnormalities), bone marrow failure, pulmonary fibrosis, and increased cancer (Armanios and Blackburn, 2012). Mutations in DC target each of the known components of telomerase and interfere with many steps in the telomerase pathway, including assembly, trafficking, recruitment to telomeres, and catalytic activity (Batista and Artandi, 2013). TCAB1 mutations occur in an autosomal recessive form of DC, in which single amino acid substitutions in TCAB1 protein cause a marked reduction in protein accumulation and disruption of telomerase localization from Cajal bodies to nucleoli (Zhong et al., 2011). Mislocalization of telomerase to nucleoli causes telomere shortening and DC; however, the mechanism by which these mutations disrupt TCAB1 function is not understood.

Nascent polypeptides must navigate a complex thermodynamic and kinetic landscape to acquire and maintain their native, functional form. Coordinating this process across all the proteins in the cell requires a highly organized proteostasis network that regulates protein biogenesis, conformational maintenance, and degradation (Kim et al., 2013; Wolff et al., 2014). For many proteins, proper folding and function requires stabilization of folding intermediates, a process performed by molecular chaperones, which derive from one of several classes, including HSP70, HSP90, and chaperonins. Responsible for the vast majority of chaperone-mediated folding within the cell, the HSP70 and HSP90 systems tend to act as monomers or homodimers to bind sections of nascent polypeptides, stabilizing regions of hydrophobicity until the polypeptides can properly fold. In contrast, chaperonins form barrel-like cages in which individual polypeptides are encapsulated, allowing folding to proceed isolated from the general cellular environment (Horwich et al., 2007).

As a general rule, fast-folding proteins are stabilized by the HSP70 and HSP90 systems, whereas difficult-to-fold proteins are transferred to the chaperonin system. TRiC (TCP-1 Ring Complex) is a large group II chaperonin complex containing eight homologous subunits arranged in two stacked, octameric rings. TRiC was originally identified based on its essential role in folding the cytoskeletal proteins actin and tubulin (Frydman et al., 1992; Gao et al., 1992). Approaches to identify novel TRiC substrates suggest that its substrates are more numerous, although only a restricted set of proteins out of the thousands that flux through the proteostasis network depend on TRiC for their biogenesis (Yam et al., 2008). The primary function of TRiC is to participate in the folding of newly synthesized polypeptides, as opposed to the refolding of stress-denatured proteins; consequently, unlike other arms of the proteostasis network, TRiC expression is linked to protein synthesis and is not induced by stress (Albanèse et al., 2006). TRiC has been shown to inhibit polyglutamine aggregation and thus may have neuroprotective properties regarding the development of age-related misfolding diseases

such as Huntington's disease (Behrends et al., 2006; Kitamura et al., 2006; Tam et al., 2006).

In this study, we deploy a genetic screen to identify regulators of telomerase trafficking using a genome-wide, high-content, RNA fluorescence in situ hybridization (FISH)-based siRNA screen in human cells. The screen identified subunits of the chaperonin TRiC, and we demonstrate an essential role for TRiC in telomerase function and trafficking of telomerase and scaRNAs to Cajal bodies. These effects are due to the essential role of TRiC in folding TCAB1, and we find that DC patient mutations in TCAB1 disrupt this critical step in TCAB1 biogenesis, providing a molecular explanation for these cases of human disease. These findings reveal important connections linking proteostasis pathways to telomerase function and telomere disease states.

RESULTS

TRiC Is a Recurrent Hit in a Genetic Screen for Regulators of Telomerase Trafficking

The application of a genetic screen interrogating the telomerase trafficking pathway has potential to identify new regulators of telomerase. However, low expression of the telomerase catalytic core, TERT and TERC, has precluded the development of such assays. To overcome these limitations, we designed and performed a genome-wide, high-content siRNA screen for genes required for localization of TERC and TCAB1 in Cajal bodies. To capture a sufficient population of cells, we acquired all images at 200 \times magnification. Although endogenous TERC foci were readily detectable at higher magnifications (i.e., 630 \times), 200 \times magnification was insufficient to resolve individual foci (Figure S1A available online); instead, we generated clonal HeLa S3 cells stably overexpressing TERC at levels \sim 9-fold higher than endogenous TERC (Figure S1B). Overexpressed TERC was readily detectable in foci at 200 \times (Figures S1A and S1B) and retained its dependence on dyskerin and TCAB1 for localization to Cajal bodies, validating the utility of this approach (Figure S1C).

Using a human siRNA library containing 21,121 siRNA pools (4 siRNAs per pool), each targeting a single gene, we reverse transfected TERC-expressing HeLa S3 cells in 384-well optical plates in triplicate (Figure 1A). Cells were fixed, and immunofluorescence was performed for coilin, a Cajal body marker, and TCAB1 using antibodies against each endogenous protein, followed by RNA FISH for TERC. Images of a single field in each well (\sim 100 cells) were captured robotically, and images were loaded into an analysis pipeline, which automatically identified nuclei, coilin foci, TCAB1 foci, and TERC foci. For each image, we calculated a localization score—the fraction of foci pixels in the TERC or TCAB1 image that were also foci in the coilin image, i.e., foci overlap using coilin as the baseline. Scores were normalized to negative controls to calculate “localization inhibition,” with 0% indicating no TERC or TCAB1 disruption and 100% representing complete lack of TERC or TCAB1 in Cajal bodies (see Figure 1B for example images). For TERC and TCAB1, the average Z'-factors across all plates were 0.68 and 0.81, respectively, well above the conventional 0.5 value that designates an assay amenable to large-scale screening (Figure S1D) (Zhang

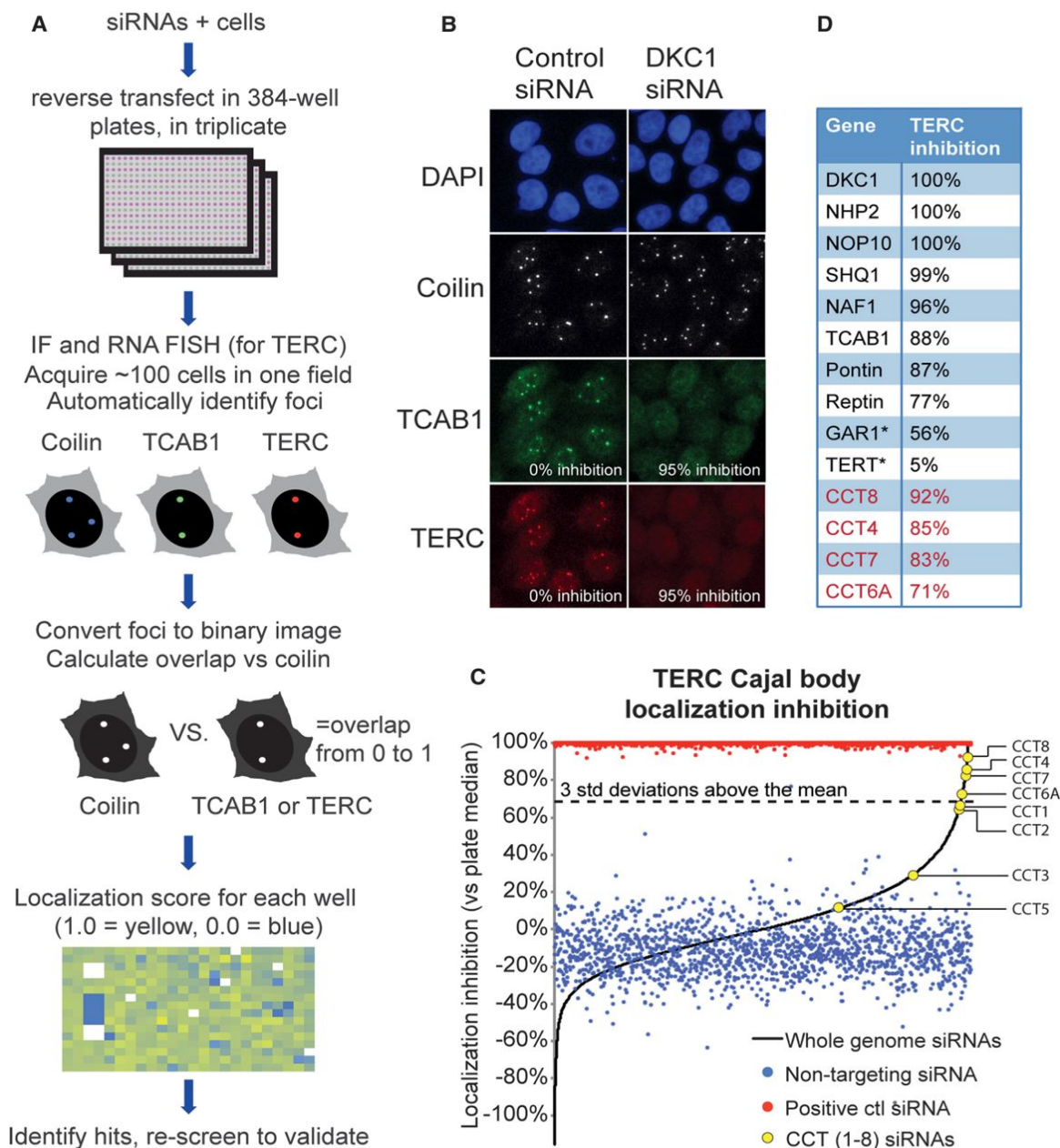


Figure 1. TRiC Is a Recurrent Hit in a Genetic Screen for Regulators of Telomerase Trafficking

(A) High-content screening workflow. Reverse transfection of HeLa cells into 384-well plates with siRNA pools targeting individual genes, stained for Coilin, TCAB1, and TERC and imaged at 200× to determine overlap between the Coilin and TERC or TCAB1 channels. Heat map of data from one plate; well color indicates localization score. Yellow, localization; blue, no localization.

(B) Representative screen images for negative control and positive control (Dyskerin siRNA) images, shown with percent localization inhibition.

(C) Graphical representation of screen data. y axis, percent inhibition of TERC localization to Cajal bodies. x axis, dimensionless. Red dots, positive control wells (siDKC1; n = 1,608). Blue dots, negative control wells (siNon-targeting; n = 1,608). Black line represents the median value (of triplicates) from each siRNA pool in ascending order (n = 21,119). Yellow dots show siRNA pools targeting CCT/TRiC genes. Dashed line, three SD cutoff.

(D) Localization inhibition for known telomerase components and biogenesis factors (black) and TRiC subunits called as hits (red). All called as hits except those marked by an asterisk.

et al., 1999). Whereas positive control dyskerin siRNA wells (eight per plate, n = 1,608) yielded a median TERC localization inhibition of 99.6%, the median TERC inhibition of all the experimental siRNAs was −0.3% with a SD of 22.4%, demonstrating

that the vast majority of experimental siRNAs had no effect on TERC localization to Cajal bodies (Figure S1C). After removing toxic siRNAs and coilin-eliminating siRNAs and applying a hit cutoff of three SDs above the mean (corresponding to 67.3%

inhibition), we found that 119 siRNA pools qualified as TERC inhibition hits (Figure 1C). Hits included virtually all known components of the H/ACA box snoRNP complex, which is required for TERC biogenesis: DKC1, NHP2, NOP10, NAF1, and SHQ1 (Figure 1D). Hits also included TCAB1 and the telomerase assembly factors Pontin and Reptin. TERT siRNAs had no effect on TERC localization, which may be a consequence of poor knockdown. In addition to these known telomerase biogenesis and assembly factors, which validate the screening approach, hits included four members of TRiC: CCT8, CCT4, CCT7, and CCT6A (Figure 1D). Two other TRiC members fell just below the hit cutoff: CCT1 and CCT2 (Figure 1C). CCT6B siRNAs had no effect, which was expected as its expression is restricted to testicular tissues (Kubota et al., 1997). In the primary screen, CCT3 and CCT5 also showed no effect, which we subsequently determined were false negatives (see Figure S2A). Analyzing the TCAB1 channel in a similar manner, we found that TRiC subunits were also required for TCAB1 localization to Cajal bodies (Figure S1F), demonstrating that multiple telomerase components were perturbed due to TRiC depletion.

TRiC Is Required for TERC and scaRNA Trafficking

To validate and expand these results at a conventional scale, we depleted TRiC components in HeLa cells using RNA interference and assessed TERC localization at high magnification. Depletion of TRiC subunits led to mislocalization of endogenous TERC to nucleoli, results indistinguishable from depletion of TCAB1 itself (Figures 2A and 2D) ($p < 0.001$ by Fisher's exact test). In contrast, depletion of dyskerin (DKC1) resulted in a global reduction in TERC signal, which is consistent with the function of dyskerin in TERC stability and assembly. Unlike depletion of dyskerin, neither TCAB1 depletion nor depletion of any TRiC subunit reduced TERC level measured by northern blot (Figure 2B). Localization of TERC stably overexpressed in HeLa cells was similarly dependent on TRiC, as depletion of any TRiC subunit led to mislocalization to nucleoli, demonstrating that overexpression of TERC cannot compensate for TRiC loss (Figure S2A). To test whether TRiC was required more generally for localization of scaRNAs in Cajal bodies, we used FISH to examine trafficking of scaRNAs containing either an H/ACA box only (U93) or an H/ACA box and a C/D box (U85 and U87). In all cases, we found that TRiC knockdown resulted in partial mislocalization of these scaRNAs to nucleoli to a degree that phenocopied TCAB1 knockdown (Figures 2C, 2D, S2B, and S2C) ($p < 0.001$ by Fisher's exact test). These data demonstrate that TRiC is required for the proper localization of TERC, as well as multiple scaRNAs, to Cajal bodies.

Telomerase Recruitment and Telomere Addition Depend upon TRiC

Association between TERC and TCAB1 is required for telomerase to localize to Cajal bodies and, subsequently, for telomerase to localize to telomeres (Cristofari et al., 2007; Stern et al., 2012; Venteicher et al., 2009; Zhong et al., 2011, 2012). To assess telomerase at telomeres, we overexpressed both telomerase RNA component (TERC) and hemagglutinin antigen (HA)-tagged TERT in HeLa cells, which causes ectopic telomerase to localize to telomeres in a high proportion of cells (Cristo-

fari and Lingner, 2006; Zhong et al., 2012). We visualized the colocalization of HA-TERT and telomeres by immunofluorescence using anti-HA antibodies and antibodies against the telomere-binding protein TRF2 (Figure 3A). Telomerase was efficiently recruited to telomeres in control cells, whereas siRNA-mediated depletion of DKC1 or TCAB1 strongly decreased that recruitment. Depletion of TRiC subunits significantly reduced the average frequency of telomerase foci at telomeres from 15 HA-TERT foci at telomeres in control cells to 2–4 in TRiC-depleted cells ($p < 0.0001$, Student's *t* test) (Figure 3B). Thus, recruitment of telomerase to telomeres—a rate-limiting step in telomere synthesis—is exquisitely dependent on the TRiC complex.

To determine whether telomere addition requires TRiC, we employed an assay that allows direct visualization of newly added telomere repeats, programmed by a TERC variant containing a mutant template region (Diolaiti et al., 2013). When expressed in HeLa cells for 3 days, this mutant TERC (TSQ1), together with endogenous TERT and other telomerase cofactors, efficiently mediated addition of variant GTTGCG repeats that were visualized by combined DNA FISH for the mutant sequences and for natural telomere repeats (TTAGGG) (Figure 3C). Control cells contained an average of approximately ten telomeres per cell with detectable GTTGCG signal, demonstrating that telomere elongation occurred at these telomeres over the 4 day period (Figure 3D). Depletion of TRiC subunits, DKC1, or TCAB1 using siRNA strongly decreased addition of new repeats to fewer than three telomeres per cell, indicating that TRiC is required for telomere synthesis. Together, these data show that loss of TRiC leads to mislocalization of telomerase to nucleoli, inhibits telomerase recruitment to telomeres, and prevents telomere addition by telomerase.

TRiC Interacts with the TCAB1 WD40 Domain and Is Required for TCAB1 Accumulation

The similar requirements for TCAB1 and TRiC in telomerase localization, recruitment to telomeres, and telomere synthesis, together with the observation that TRiC knockdown led to reduced TCAB1 foci in our genetic screen (Figure S1F), led us to hypothesize that the requirement for TRiC in telomerase function was based upon a role for TRiC in TCAB1 biogenesis. To test this idea, we depleted TRiC subunits in HeLa S3 cells and monitored TCAB1 foci by immunofluorescence and overall TCAB1 levels by western blot. Depletion of any of the eight TRiC subunits with siRNA strongly diminished TCAB1 foci (Figures 4A and S3A) and markedly reduced TCAB1 protein levels (Figure 4B) without altering TCAB1 mRNA levels (Figure S3B). Each CCT siRNA led to codepletion of specific subsets of CCT proteins, likely reflecting patterns of codependence of CCT components within the TRiC complex (Figure 4B). To determine the selectivity of this TRiC-TCAB1 relationship, we depleted two Hsp70 family molecular chaperones: HSPA8 (Hsc70) and HSPA5 (GRP-78), both of which are required for the proper folding of a diverse set of cytoplasmic proteins (Liu et al., 2012). Depletion of these Hsp70 family chaperones had no effect on TCAB1 protein accumulation (Figure S3C), showing that disruption of TCAB1 is a TRiC-specific response rather than a result of general protein folding disruption.

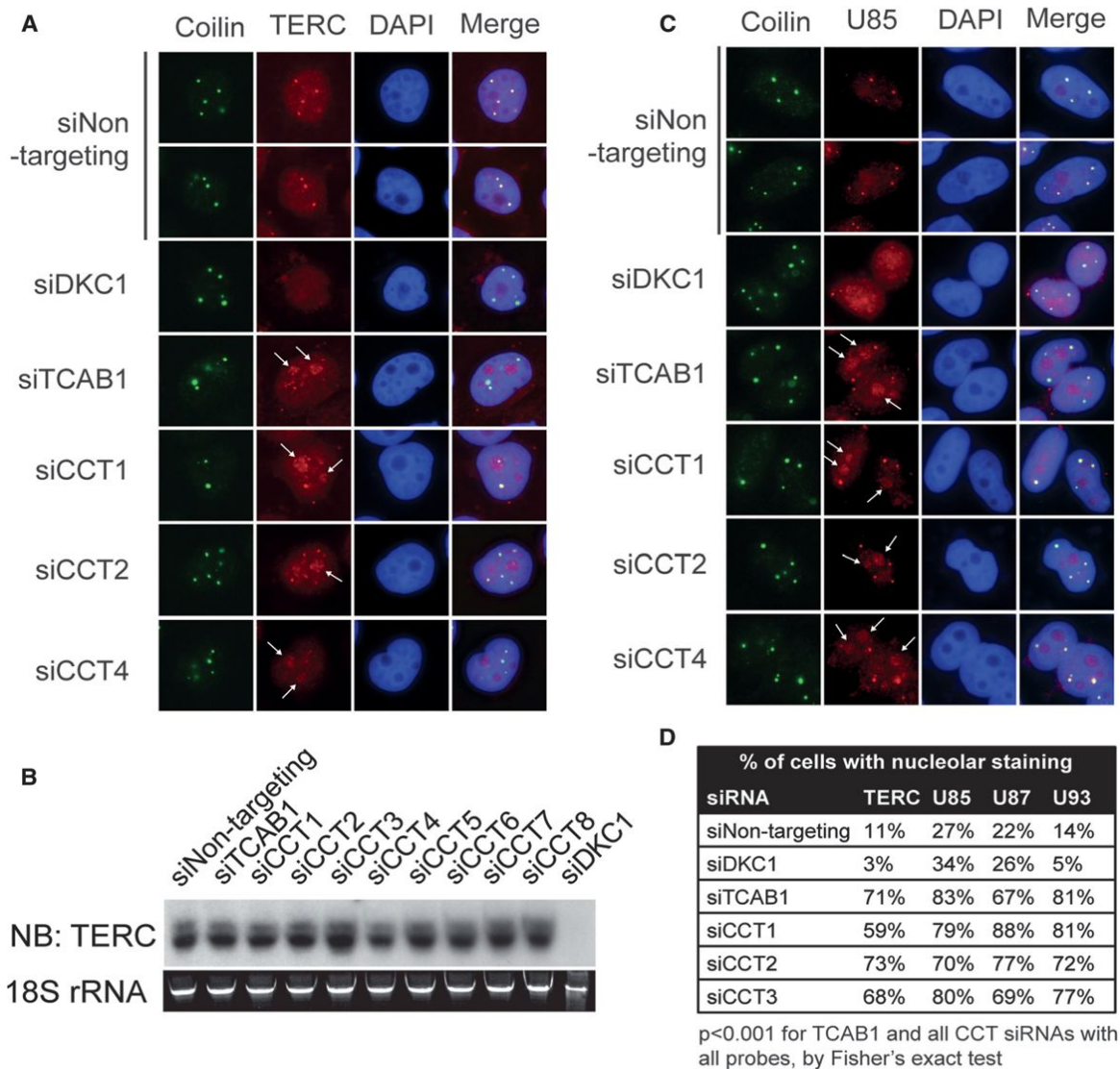


Figure 2. TriC Is Required for TERC and scaRNA Trafficking

(A) RNA FISH for TERC (red) and IF for coilin (green) in HeLaS3 cells transfected with indicated siRNAs. DAPI, blue. Arrows indicate nucleolar staining.

(B) Northern blot for TERC in HeLa S3 cells transfected with indicated siRNAs; 18S rRNA, loading control.

(C) RNA FISH for U85 (red) and IF for coilin (green) in HeLaS3 cells transfected with indicated siRNAs. DAPI, blue. Arrows indicate nucleolar staining.

(D) Quantification of nucleolar staining for TERC, U85, U87, and U93 scaRNAs from (A) and (C) and Figures S2B and S2C. 150 cells analyzed per condition, p < 0.001 for either TCAB1 siRNAs or each CCT siRNAs versus siNon-targeting controls, Fisher's exact test.

To determine whether TCAB1 and TriC interact, we purified TCAB1 complexes through dual-affinity chromatography coupled with mass spectrometry (MS). TCAB1 tagged at its N terminus with Staph Protein A, a TEV cleavage site, and three HA epitopes (AH3-TCAB1) was stably expressed in HeLa S3 cells by retroviral transduction, purified on rabbit IgG, eluted with TEV protease, captured again using anti-HA resin, and then eluted, fractionated by SDS-PAGE, and analyzed using nano-liquid chromatography-tandem MS (nanoLC-MS/MS). High peptide coverage was obtained for TCAB1, for the TCAB1-associated protein dyskerin, and for all eight subunits of the TriC complex (Figure 4C). Immunoprecipitation of endog-

enous TCAB1 using TCAB1 polyclonal antibodies effectively depleted HeLa cell extracts of TCAB1 protein and revealed endogenous CCT1 and CCT2 in association with TCAB1 (Figure 4D). Pull-down of TCAB1 did not deplete CCT1 or CCT2, indicating that only a small fraction of TriC is associated with TCAB1 (Figure 4D). These findings show that TCAB1 and TriC associate at the endogenous level and suggest a substoichiometric or transient association.

The predicted structure of TCAB1 includes an unstructured N-terminal region, a WD40 domain with seven WD40 repeats forming a β propeller—a protein structure with complex topology and multiple regions of hydrophobicity—followed by a short

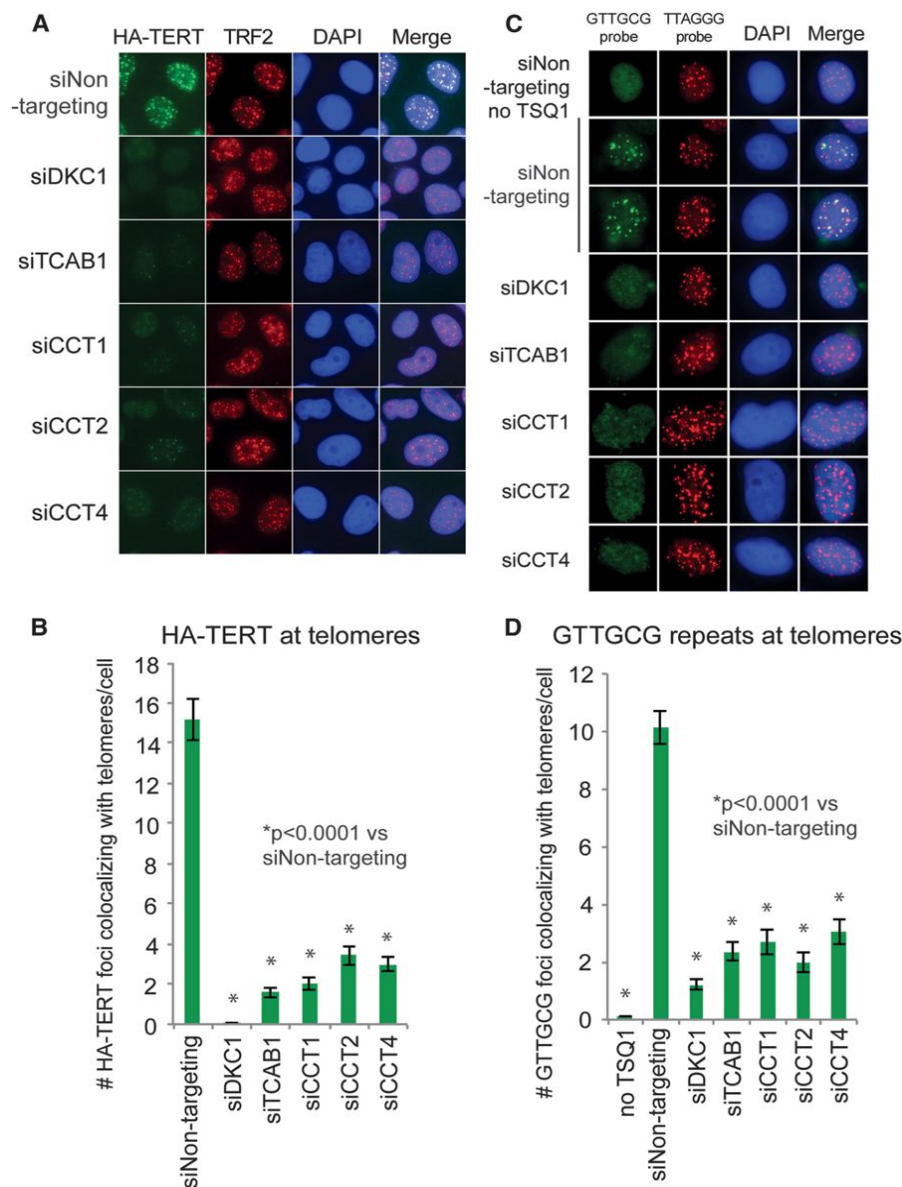


Figure 3. Telomerase Recruitment and Telomere Addition Depend upon TRiC

(A) Colocalization of HA-TERT and TRF2 by IF in cells transfected with indicated siRNAs and HA-TERT and TERC plasmids.

(B) Quantification of HA-TERT at telomeres from (A). 150 cells analyzed per condition. Mean \pm SEM, $p < 0.0001$ for each siRNA versus siNon-targeting control by two-tailed Student's t test.

(C) FISH for incorporation of mutant telomeres (GTTGCG) and FISH for wild-type telomere sequences (TTAGGG) in HeLa S3 cells transfected with indicated siRNAs and mutant TERC (TSQ1).

(D) Quantification of GTTGCG foci at telomeres (TTAGGG foci) from (C). 150 cells analyzed per condition. Mean \pm SEM, $p < 0.0001$ for each siRNA versus siNon-targeting control by two-tailed Student's t test.

to TRiC compared to full-length TCAB1. These data show that the WD40 domain of TCAB1 mediates association with TRiC but that stabilization of TCAB1 and efficient release from TRiC also require the C-terminal region of TCAB1.

To determine which TCAB1 sequences are involved in binding TERC, we performed IP-northern assays in RRL with FLAG-tagged TCAB1 mutant proteins after adding in-vitro-transcribed TERC. In validation studies, the association between wild-type TCAB1 and TERC in this assay was dependent upon an intact CAB box, as either of two inactivating mutations in the CAB box, m1-UGAc or m2-UcAG (Cristofari et al., 2007; Jády et al., 2004), markedly diminished binding to TCAB1 (Figure S3F). Full-length TCAB1 and Δ N125 bound wild-type TERC similarly well, but association between Δ C and WD40 fragments of TCAB1 with TERC was greatly reduced (Figure 4E). These data suggest that the

WD40 domain, together with the short C-terminal extension, may comprise a single functional domain that both binds TERC and depends upon TRiC for folding.

TRiC Is Required to Fold TCAB1, Enabling TCAB1 to Bind TERC

Although the kinetics of dissociation vary by substrate, TRiC clients are only transiently associated with the complex and are released after folding is complete. To determine whether TCAB1 is a TRiC substrate, we in vitro translated TCAB1 or yeast actin as a positive control in the presence of 35 S-methionine in RRL. After the initial labeling, reactions were chased with cold methionine for the indicated number of hours, and proteins were fractionated by native PAGE. Actin is a known substrate of TRiC and is folded in a TRiC-dependent manner in RRL

C-terminal extension (Figure 4E). Analysis of a series of N-terminal deletion mutants of TCAB1 revealed that the WD40 domain and the C-terminal extension were sufficient for TRiC association and for Cajal body localization in HeLa cells (Figures S3D and S3E). Further impingement on the WD40 domain or deletion of the short C-terminal extension prevented accumulation in human cells (data not shown); therefore, we expressed additional mutant proteins in rabbit reticulocyte lysate (RRL), an extract that enables translation and folding but lacks active protein degradation machinery. Deletion of the C-terminal region (Δ C, deletion of residues 510–548) resulted in efficient translation and an enhanced association with rabbit TRiC in the extract by coimmunoprecipitation (co-IP) (Figure 4E). Expression of the WD40 domain alone (residues 126–509, WD40) resulted in low levels of this isolated domain but in significantly greater binding

(Gao et al., 1992). Migration of newly synthesized yeast actin was retarded in the gel, whereas after a 1 hr chase, actin migrated more rapidly, which is consistent with being released from TRiC (Vainberg et al., 1998) (Figure 5A). In-vitro-translated TCAB1 similarly migrated in two forms. Newly translated TCAB1 protein was detected almost exclusively in a slow migrating species that comigrated with newly synthesized actin. Over a 4 hr chase period, TCAB1 became rapidly migrating and formed a discrete lower band. We verified that both radiolabeled species were in fact TCAB1 by translating FLAG-TCAB1 and performing a western blot after native PAGE (Figure S4A). TRiC, visualized by CCT2 western blotting, comigrated with the retarded, newly synthesized form of TCAB1 in RRL (Figure S4A). We showed that the upper TCAB1 band was TRiC bound by immunodepleting CCT2 after TCAB1 translation (Figure S4B), leading to a significant decrease in the upper bands of both yeast actin and TCAB1 while preserving the lower bands for both proteins (Figure 5B). These results demonstrate that TCAB1 polypeptides are bound by TRiC during or soon after translation and are released from association with TRiC over time.

We hypothesized that TRiC released TCAB1 once folding was complete. Partial resistance to protease digestion is a characteristic of many correctly folded, compact proteins (Frydman et al., 1994). To optimize this approach for natively folded TCAB1, we performed a dose response proteinase K (PK) digestion on HeLa lysate and analyzed protease-resistant TCAB1 fragments by western blotting (Figure S4C). Partial PK digestion at 1.2 μ g/ml generated two protease-resistant TCAB1 bands between 37 and 50 kDa, whereas higher concentrations degraded the protein completely and lower concentrations had no effect. We then examined the protease resistance of in-vitro-translated TCAB1 before and after TRiC-release to assess whether TRiC binding caused TCAB1 to adopt a similarly protease-resistant structure. PK degraded all protein at zero hr but resulted in a protected species at 4 hr, characterized by two discrete protease-resistant TCAB1 bands (Figure 5C) similar in size to those generated by partial digestion of cellular TCAB1 (Figure S4C). Importantly, the input of total TCAB1 protein treated with PK was comparable at both time points (Figure 5C, bottom). These data indicate that TCAB1 is partially resistant to protease treatment only after being released from TRiC, suggesting that this TRiC-released TCAB1 species is compact and correctly folded.

To understand whether TRiC is required for folding TCAB1, we immunodepleted TRiC from RRL using antibodies to CCT2 and then reconstituted the complex to endogenous levels using purified bovine TRiC or used mock depletion with IgG as a negative control (Figure 5D). Proteins were then synthesized under TRiC-depleted and TRiC-add-back conditions. Translation of either yeast actin or TCAB1 in TRiC-depleted RRL resulted in a marked reduction of both the TRiC-associated form (>50% reduction versus mock-depleted controls) and the folded species for each protein (>90% reduction versus mock-depleted controls) by native PAGE (Figures 5E and S4D). Add-back of bovine TRiC to the depleted RRL restored both the TRiC-bound (>100% versus mock-depleted controls) and folded species (>50% versus mock-depleted controls) of actin and TCAB1 (Figures 5E and S4D). Depletion of TRiC, or TRiC add-back, did not

significantly change the overall amount of translated proteins by SDS-PAGE (Figure 5E, bottom). These data show that TRiC is required for TCAB1 to be folded to a compact, rapidly migrating conformation in RRL.

TCAB1 associates with TERC and with scaRNAs by binding the CAB box sequence common to these RNAs (Tycowski et al., 2009; Venteicher et al., 2009). To understand whether TRiC-mediated folding is required for TCAB1 to associate with TERC, we tested the effects of TRiC depletion in RRL on TCAB1-TERC binding by IP-northern (Figure 4). TCAB1 translated in TRiC-depleted RRL bound <10% of the TERC bound by TCAB1 translated in mock-depleted RRL (Figures 5F and S4E). Add-back of bovine TRiC to TRiC-depleted extracts restored the ability of TCAB1 to associate with TERC to >60% of mock-depleted level (Figures 5F and S4E). Taken together, these data demonstrate that TCAB1 is an obligate TRiC substrate, requiring TRiC for folding and TERC binding, and that loss of TRiC cannot be compensated by the presence of other chaperones.

Disruption of TRiC-Mediated Folding Underlies the Defects in the TCAB1 Mutant Form of DC

The effects of TRiC depletion on the telomerase pathway—diminished TCAB1 protein, loss of TERC and TCAB1 from Cajal bodies, and a failure of telomere synthesis—closely resembled the underlying features of a form of DC caused by compound heterozygous mutations in TCAB1 (Zhong et al., 2011). For each of four identified disease-causing alleles, missense mutations in TCAB1 were located in or near loop regions within the WD40 domain, the principal substrate for TRiC-mediated folding (Figure 6A). We therefore hypothesized that the single amino acid mutations in TCAB1 specifically disrupted folding of TCAB1 by TRiC. TCAB1 protein levels were markedly reduced in Epstein-Barr virus (EBV)-transformed lymphoblasts derived from two DC patients with compound heterozygous TCAB1 mutations, compared to levels in unrelated individuals with wild-type TCAB1 (Figure 6B). Furthermore, each of the mutant TCAB1 alleles also showed impaired protein accumulation when stably overexpressed in HeLa cells, indicating that diminished expression is an intrinsic property of the mutant proteins and is not related to mRNA level (Figures 6B and S5A). Consistent with diminished expression, TCAB1 protein failed to localize properly to Cajal bodies, both in patient lymphoblasts (Figure 6C) and in stably transduced HeLa cells in which TCAB1 was detected in the cytoplasm (Figure S5B), where TRiC resides almost exclusively (Figure S5C) (Zhong et al., 2011).

To determine whether DC-derived TCAB1 mutants fail to accumulate because of a defect in folding by TRiC, we assessed the resistance to protease digestion of mutant versus wild-type TCAB1 in patient-derived lymphoblasts. As in HeLa cells and RRL, wild-type TCAB1 in lymphoblasts displayed two protease-resistant bands between 50 and 37 kDa; however, mutant TCAB1 in DC-patient-derived lymphoblasts showed no resistance to protease treatment under the same conditions (Figure S5D). Mutant TCAB1 accumulated to lower levels than wild-type TCAB1, raising the possibility that the difference in protease resistance was due to a difference in total protein level. We therefore examined DC-derived TCAB1 mutants versus

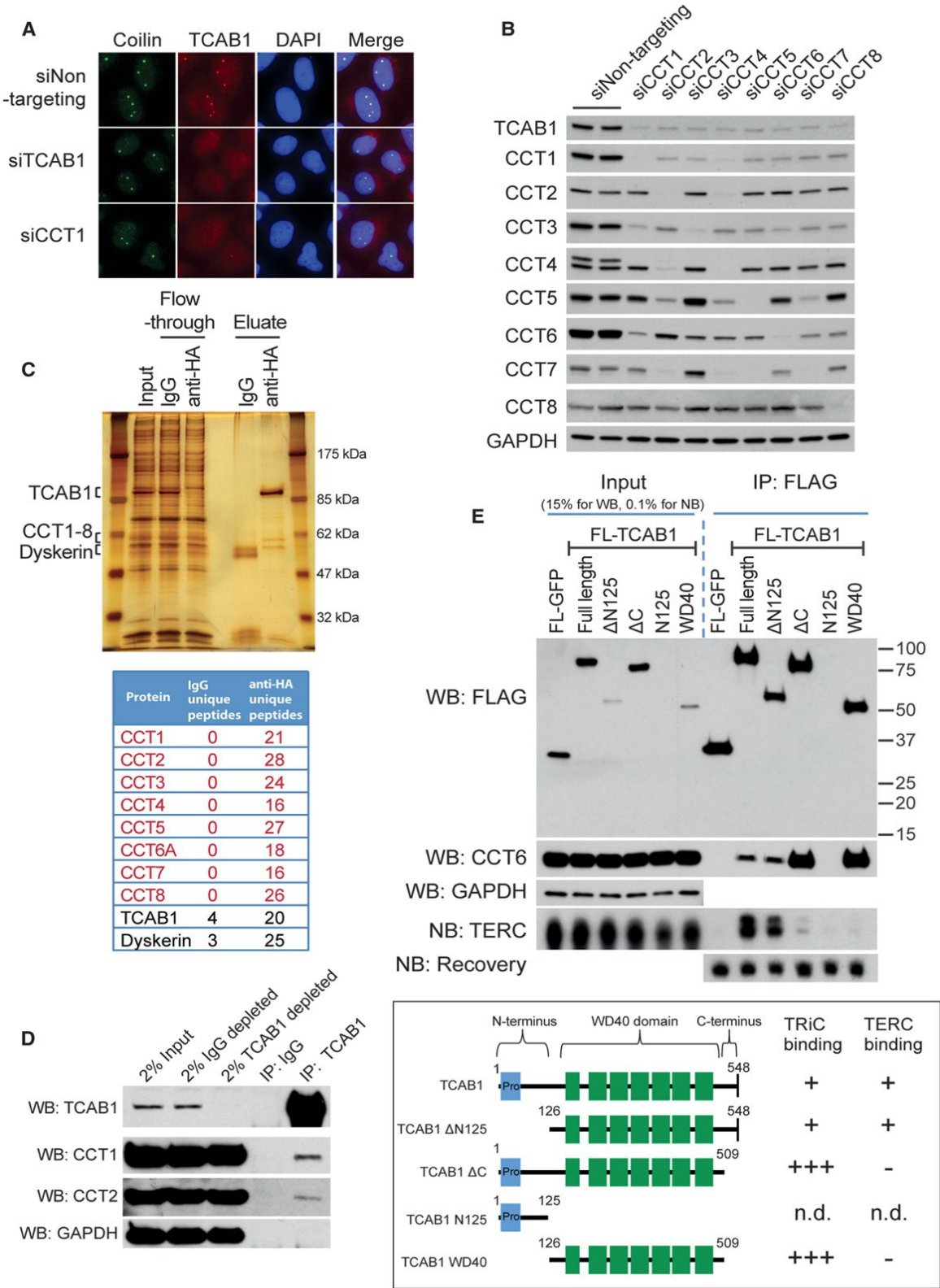


Figure 4. TRiC Interacts with the TCAB1 WD40 Domain and Is Required for TCAB1 Accumulation
(A) IF for TCAB1 (red) and coilin (green) in HeLa S3 cells transfected with siRNAs against TCAB1 and CCT1. DAPI, blue.
(B) Western blot for TCAB1 and TRiC subunits in HeLa S3 cells transfected with siRNAs against eight TRiC subunits; GAPDH, loading control.
(legend continued on next page)

wild-type TCAB1 expressed in RRL, where the mutant and wild-type TCAB1 proteins accumulate at comparable levels (Figure 6D, bottom). Unlike wild-type TCAB1, which was partially protected from PK digestion at 4 hr, each DC-derived TCAB1 mutant failed to show PK resistance at 4 hr (Figure 6D, top). To determine whether this inability to fold properly corresponded to differences in TRiC binding or release, TCAB1 proteins were assayed after pulse-chase in RRL by native PAGE. In-vitro-translated TCAB1 mutants were initially bound by TRiC to the same degree as wild-type TCAB1, but whereas wild-type TCAB1 was released after several hours, the TCAB1 mutants remained TRiC associated and were impaired in progression to the compact, rapidly migrating form (Figure 6E; see Figure 6D for loading). We assessed whether this impaired folding of the DC-derived TCAB1 mutants affected their ability to bind TERC. Compared with wild-type TCAB1, in-vitro-translated TCAB1 mutants showed a marked decrease in TERC association by IP-northern (Figure 6F). These data show that DC-associated TCAB1 mutants cannot be properly folded by TRiC, show enhanced TRiC association, and are impaired in their ability to bind TERC.

The increased association of TCAB1 mutants with TRiC was recapitulated in human cells, as measured by IP-western of HA-tagged TCAB1. Despite accumulating at a lower level, each TCAB1 mutant showed a steady-state association with TRiC substantially greater than that of wild-type TCAB1 (Figure 6G). We used extracts from DC patient lymphoblasts to determine whether this increased association with TRiC was evident when TCAB1 mutants were expressed at endogenous levels. Although TCAB1 mutant proteins were expressed at reduced levels in DC lymphoblasts, immunoprecipitation of endogenous mutant TCAB1 protein pulled down significantly more TRiC compared with immunoprecipitation of wild-type TCAB1 from healthy donor control lymphoblasts (Figure 6H). Taken together, these data show that the TCAB1 mutant form of DC is caused by amino acid mutations that disrupt TCAB1 folding by TRiC, providing important evidence that telomere maintenance by telomerase is exquisitely dependent on TRiC and that interfering with this process results in severe disease in humans.

DISCUSSION

A Model for TRiC-Mediated Folding of TCAB1 in Telomere Maintenance

In searching for regulators of telomerase trafficking using a genome-wide loss-of-function high-content screen, we identified the TRiC chaperonin as a critical node in the telomerase pathway and one whose actions are subverted in disease. Our data support a model in which newly translated TCAB1 protein associates with TRiC through an interaction that requires the

WD40 domain of TCAB1. TRiC folds TCAB1 in the cytoplasm, and once correctly folded, TCAB1 translocates to the nucleus where it binds the TERC CAB box and likely associates with dyskerin through protein-protein contacts to facilitate localization of telomerase within Cajal bodies. This localization step is essential for telomere maintenance, as TCAB1 is also required for recruitment of telomerase to telomeres, a step governed by the interaction between TERT and TPP1 (Nandakumar et al., 2012; Stern et al., 2012; Venteicher et al., 2009; Zhong et al., 2011, 2012). Consistent with a requirement in folding TCAB1, depletion of any TRiC subunit resulted in reduced TCAB1 protein in Cajal bodies, mislocalization of TERC to nucleoli, and a marked impairment of telomerase recruitment to telomeres. TRiC is also needed for correct trafficking of scaRNAs to Cajal bodies, as loss of TRiC resulted in inappropriate trafficking of scaRNAs to nucleoli. Because scaRNAs serve as guides for the pseudouridylation of splicing RNAs, our data suggest that, through folding TCAB1, TRiC may be required for the assembly of splicing RNPs and perhaps for efficient RNA splicing.

Our data indicate that TRiC-dependent folding of TCAB1 represents a third biogenesis pathway required for telomerase function. Most well understood are the RNA-dependent steps needed for stability, processing, and assembly of TERC. TERC biogenesis relies on the machinery devoted to related snoRNAs, including the dyskerin core complex, assembly factors Shq1 and Naf1, and factors needed for RNA end processing and capping (Figure 7, green box). Factors have also been identified that are required for TERT accumulation and assembly into the telomerase complex. Pontin and Reptin are AAA+ ATPases that are required for both TERC accumulation and TERT assembly into the telomerase holoenzyme (Venteicher et al., 2008). Additionally, the molecular chaperones p23 and Hsp90 both bind TERT, and inhibition of this interaction reduces accumulation of TERT protein and blocks assembly of the telomerase catalytic core (Holt et al., 1999) (Figure 7, orange box). In contrast to these two pathways focused on the catalytic core of the enzyme, our findings reveal a proteostasis pathway that is not essential for the biogenesis of TERT or TERC but is critical for the folding of TCAB1 and therefore for localization of the telomerase complex in Cajal bodies (Figure 7, blue box).

TRiC-Mediated Folding of TCAB1 Is Required for Prevention of Dyskeratosis Congenita

TRiC depletion using siRNA caused biochemical and cell biological phenotypes similar to those seen in cells from DC patients with missense mutations in TCAB1, including reduced TCAB1 protein level and mislocalization of TERC to nucleoli. Based on these similarities, we investigated the hypothesis that the missense mutations in TCAB1, located in the WD40 repeat domain, represent folding mutations—sequence

(C) Top, dual-affinity chromatography purified AH3-TCAB1 from HeLa S3 cells, silver stain. Bottom, unique peptides for each protein identified by MS. IgG in eluate, pre-clear negative control step before anti-HA pull-down.

(D) Western blot for TCAB1 and TRiC subunits from co-IP of endogenous TCAB1 from HeLa S3 cells. IgG, negative control.

(E) Co-IP of in-vitro-translated FLAG constructs in RRL incubated with in vitro transcribed TERC. Top, western blot (FLAG, CCT6, and GAPDH) and northern blots (TERC and recovery control). Bottom, schematic of each construct and degree of TRiC and TERC interaction. n.d., not determined because N125 did not accumulate to detectable levels.

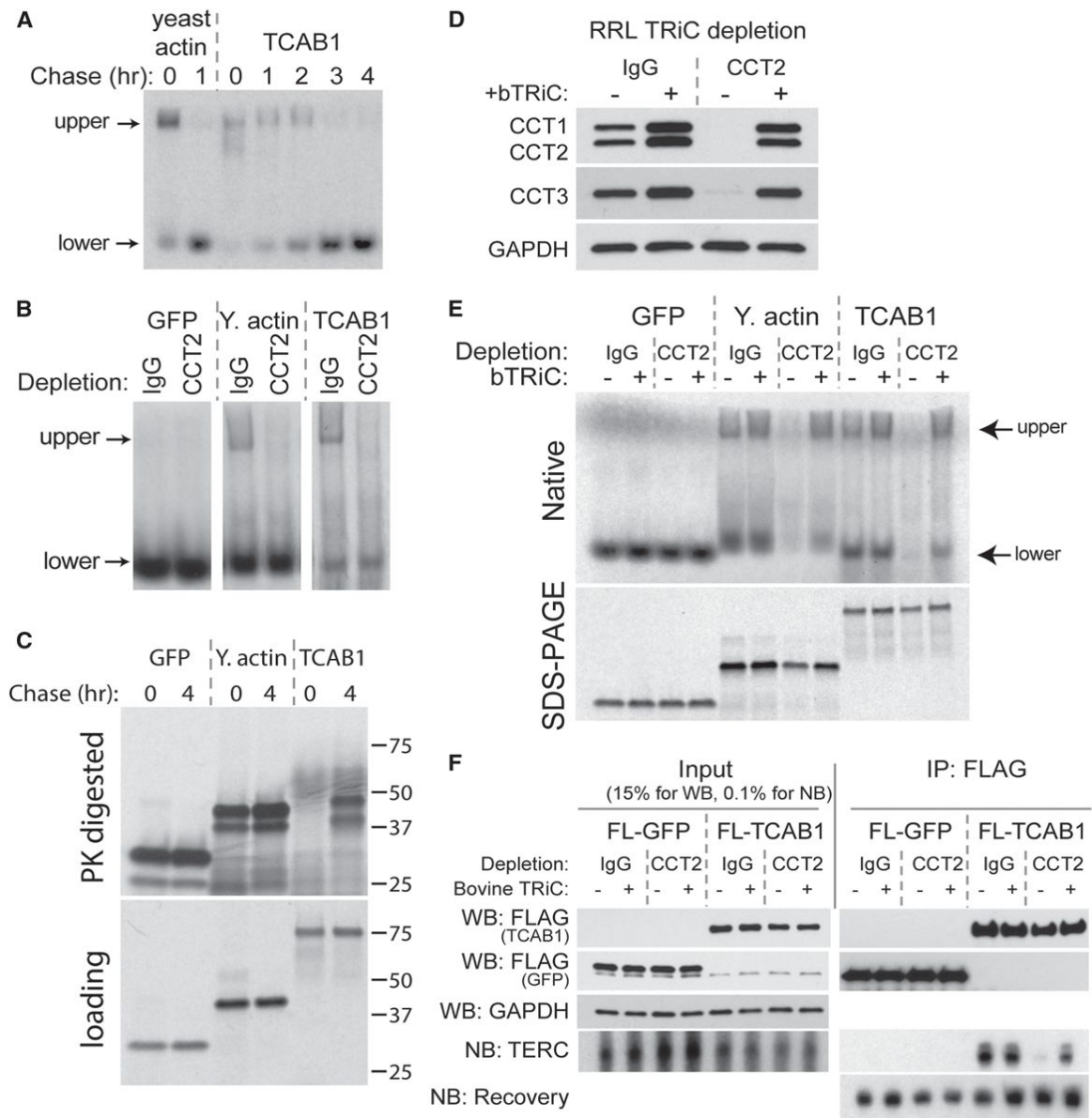


Figure 5. TRiC Is Required to Fold TCAB1, Enabling TCAB1 to Bind TERC

(A) Time course of [35 S]-Met-labeled actin or TCAB1 in RRL, chased with cold Met for indicated time, native PAGE. Upper, slowly migrating species. Lower, rapidly migrating species.

(B) Immunodepletion of CCT2 from RRL extracts in which [35 S]-Met-labeled proteins (GFP, actin, or TCAB1) were translated, native PAGE. IgG, negative control.

(C) Partial PK digestion of [35 S]-Met-labeled proteins (GFP, actin, or TCAB1) in RRL chased with cold Met for indicated times, top. Total, undigested protein, bottom. SDS-PAGE, top and bottom.

(D) Immunodepletion of CCT2 from RRL extracts. IgG, negative control. +bTRiC, add-back of purified bovine TRiC. Western blot for CCT1, CCT2, CCT3, and GAPDH.

(E) Immunodepletion of CCT2 from RRL extracts (as in D) prior to translation of [35 S]-Met-labeled proteins (GFP, actin, or TCAB1). Native PAGE, top. SDS-PAGE for loading, bottom. +bTRiC, add-back of purified bovine TRiC.

(F) Co-IP of TCAB1 and TERC from RRL. Western (FLAG, GAPDH) and northern (TERC, recovery) blots from co-IP of FLAG constructs in vitro translated in previously IgG or CCT2-immunodepleted RRL, as in (D), incubated with in-vitro-transcribed TERC. Blots were processed simultaneously and are directly comparable.

changes that disrupt the ability of TCAB1 to be correctly folded by TRiC. Our data indicated that each patient-derived TCAB1 mutant protein failed to be released from TRiC, could not

achieve a protease-resistant form, exhibited marked reduction in binding TERC, and showed a significant increase in association with TRiC. All these findings are consistent with a disruption

of TCAB1 folding by the amino acid mutations in DC patients. Because mutant TCAB1 readily binds TRiC, we conclude that the missense mutations disrupt the folding process itself and that misfolded TCAB1 is eventually released and destroyed, after failing an important quality control step. Although misfolded TCAB1 is reduced in level and is functionally impaired, it is important to note that reduced levels of mutant TCAB1 protein remain detectable in the nucleus and even in Cajal bodies. These findings suggest that the folding mutations are hypomorphic in nature, severely impairing, but not eliminating, TCAB1 function.

Folding Mutations Yield Hypomorphic Disease Variants in Highly Conserved Proteins

Thus far, no complete loss-of-function mutations in any telomerase component have been identified in DC patients, suggesting that complete inactivation of telomerase function may be incompatible with viability in humans (Zaug et al., 2013). Homozygous null mutations in TCAB1 would be even less likely, given the potential importance of TCAB1 in the pseudouridylation of splicing RNAs. The challenge in generating a telomere-disease phenotype lies in reducing telomere maintenance levels below a certain threshold, while preserving sufficient reserves in telomere maintenance to avoid embryonic lethality. For multifunctional proteins, their functions in other pathways need to be preserved as well. The best example is dyskerin, which is required for pseudouridylation of ribosomal RNAs and splicing RNAs. To avoid the embryonic lethality of dyskerin inactivation, DC mutations in dyskerin are missense mutations that cluster at a particular surface of the protein, sufficiently preserving its pseudouridylation function to allow viability, while at the same time compromising its ability to support telomerase biogenesis (Mitchell et al., 1999; Rashid et al., 2006; Wong and Collins, 2006). The exquisite dependence of TCAB1 on the TRiC pathway allowed for the emergence of TCAB1 folding mutations that accomplish this titration of TCAB1 activity by a different means—by disrupting TRiC-mediated folding and by globally reducing TCAB1 levels and function. It is interesting to note that the vast majority of disease-associated mutations in TERT and TERC are point mutations, which commonly preserve some residual activity. It is likely that a significant subset of these mutations disrupt folding of TERT protein or TERC RNA and that such mutations would operate in similar fashion to the TCAB1 mutations in reducing the level and activity of TERT or TERC.

The mechanisms by which mutations in other TRiC substrates cause disease differ. Mutant huntingtin (Htt) protein generates large aggregates, eventually overwhelming the proteostasis network and causing proteotoxicity (Behrends et al., 2006; Kitamura et al., 2006; Tam et al., 2006). Instead of aggregating, VHL mutants either prevent TRiC interaction or prevent TRiC release (Feldman et al., 2003), the latter mechanism being similar to the effects of the TCAB1 mutations. The identification of TCAB1 as a TRiC substrate is surprising because TCAB1 is both nuclear and an RNA-binding protein, which are rare among TRiC substrates (Yam et al., 2008). Thus, TRiC may be responsible for folding a larger variety of protein classes than previously thought, expanding its potential contribution to the proteostasis network.

Proteostasis, Aging, Telomerase Function, and Disease

A dramatic decline in proteostasis has been linked to aging in diverse organisms, and this decline drives age-related pathologies (Hartl et al., 2011; Taylor and Dillin, 2011). TRiC, specifically, has been implicated in human age-related degeneration via its interaction with the htt protein (Kitamura et al., 2006; Tam et al., 2006). The evidence that chaperones (TRiC, Hsp90, and p23) impinge on multiple steps of telomerase biogenesis leads us to speculate that reduced proteostasis with aging may lead to a decline in telomerase function, particularly because proteostatic networks tend to lack significant excess capacity (Morimoto, 2008). Impaired telomerase function, in turn, could lead to age-related telomere dysfunction, particularly in those tissues that depend on telomerase during aging, such as bone marrow, gastrointestinal epithelium, skin, liver, and lung.

Our data identify an exquisite dependence of telomerase on TRiC, revealing that TRiC-mediated folding of TCAB1 is essential for telomere maintenance in human tissues and that disruption of this process leads to dyskeratosis congenita. These findings have important implications in understanding telomere dysfunction in aging and in human disease.

EXPERIMENTAL PROCEDURES

Cell Culture, Transfections, and Viral Transductions

HeLa S3 and 293T cells were cultured in DMEM supplemented with 10% FBS. Lymphoblasts were cultured in RPMI 1640 supplemented with 10% FBS. Lipofectamine 2000 (Life Technologies, 11668019) was used for all cDNA transfection experiments. Dharmafect 4 (Thermo, T-2004-01) was used for all siRNA transfections. All siRNAs were siGENOME pools purchased from Dharmacon/Thermo and used at 25 or 50 nM. Cells were assayed 72 hr after transfection. For transduction, 293T cells were transfected with the appropriate target plasmid and packaging constructs (lentiviral or retroviral) overnight; 48 hr later, viral supernatant was collected. Cells were transduced in the presence of 5 μ g/ml Polybrene and selected for 3–7 days depending on the selection marker.

Western and Northern Blotting

Cells were lysed in NP40 buffer (25 mM HEPES-KOH, 150 mM KCl, 1.5 mM MgCl₂, 0.5% NP40, 10% Glycerol [pH 7.5]) supplemented with protease inhibitors (Roche, 11836153001) for 20 min on ice. Lysate was clarified by centrifugation at 16,000 $\times g$ for 10 min and quantified by BCA assay (Pierce, 23225). For western blotting, 10–50 μ g was separated by SDS-PAGE, transferred onto PVDF membrane (GE Healthcare, RPN303F), and blotted according to standard protocols. 5% milk in TBST (0.1% Tween) was used for all blocking and antibody incubation steps; TBST (0.1% Tween) was used for all washes. For northern blotting, RNA was extracted with TRIzol (Life Technologies, 15596-026), separated using 5% urea TBE gel electrophoresis, transferred onto Hybond-N membrane (GE Healthcare, RPN303N), and hybridized with ULTRAhyb (Ambion, AM8670). TERC probe was generated by random body labeling with full-length TERC cDNA.

TERC In Vitro Binding Assays

TERC or TERC mutant RNA (m1, m2) was in vitro transcribed (IVT) using the MEGAScript T7 kit (Life Technologies, AM1334). Target DNA was PCR amplified with a minimal T7 promoter at the 5' end, and 240 ng of each PCR product was used in the transcription reaction. The reaction was incubated for 6 hr at 37°C, and RNA was purified using TRIzol reagent (Life Technologies, 15596-026). In vitro transcription/translation of TCAB1 or TCAB1 mutants is described in detail in the "In Vitro Transcription/Translation" section; RRL reaction was incubated for 2 hr at 30°C, and then 1 μ g IVT RNA was added and the reaction was incubated for another 2 hr at 30°C. The reaction was diluted in

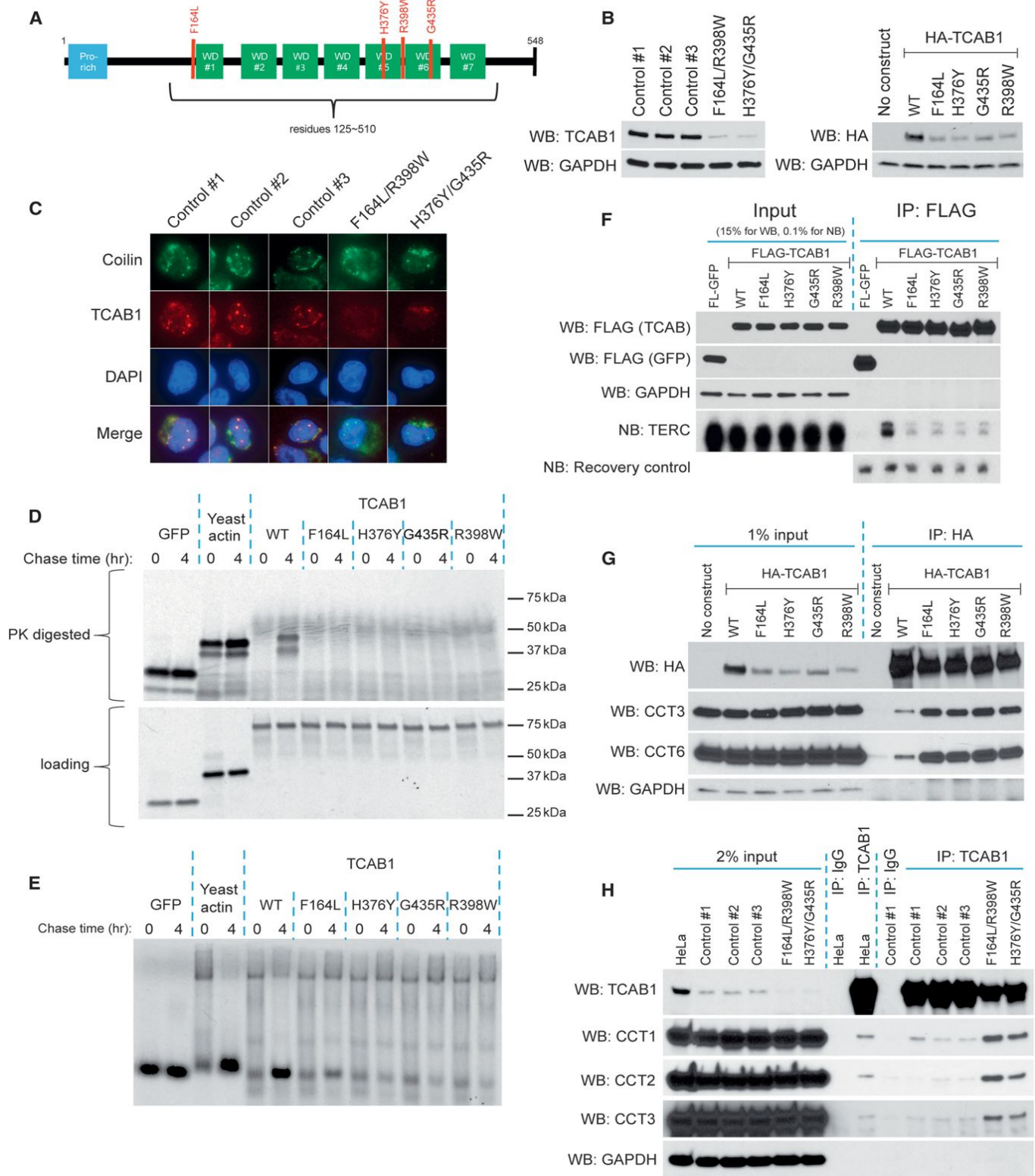


Figure 6. Disruption of TRiC-Mediated Folding Underlies the Defects in the TCAB1 Mutant Form of DC

(A) Schematic of TCAB1 domains and DC-associated mutations.

(B) Western blots of endogenous TCAB1 from control or DC patient-derived lymphoblasts (left) or stably overexpressed HA-tagged TCAB1 (wild-type or individual mutants) from HeLa S3 cells (right).

(legend continued on next page)

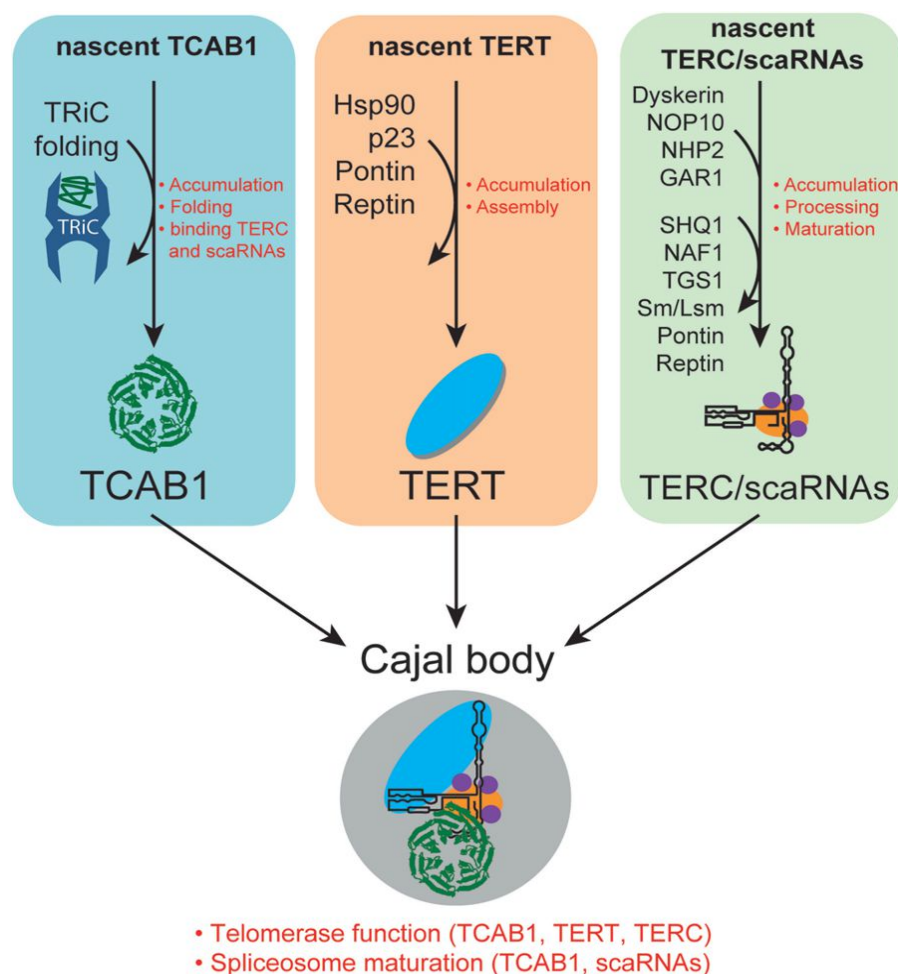


Figure 7. Model of Telomerase Biogenesis and Trafficking

Three arms of telomerase holoenzyme biogenesis (TCAB1, TERT, and TERC) undergo independent and essential biogenesis steps before being assembled into active telomerase. Curved arrows, factors transiently associated. Curved lines, factors permanently associated. Red text, processes mediated by the listed factors. Factors are not necessarily listed in chronological order and are not all inclusive.

500 μ l of NP40 buffer (see “Western Blotting” section), and FLAG immunoprecipitation was performed as described in “Immunoprecipitations and Immunodepletions.” RNA was analyzed by northern blotting.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures and five figures and can be found online at <http://dx.doi.org/10.1016/j.cell.2014.10.059>.

AUTHOR CONTRIBUTIONS

A.F., F.Z., J.F., and S.E.A. designed the experiments. A.F. and F.L.Z. performed the experiments. A.S.V., Z.M., and T.D.V. performed TCAB1 IP-MS

studies. A.F. and S.E.A. wrote the manuscript with input from all other co-authors.

ACKNOWLEDGMENTS

We thank members of the Artandi and Frydman labs for helpful discussions and advice. We thank David Solow-Cordero in the Stanford High-Throughput Bioscience Center, with support from NIH NCRR Instrumentation grant S10RR026338. We are grateful to Anthony Tomlinson for generously providing purified bovine TRiC and to Sharon Savage for supplying DC patient lymphoblasts. A.F. was supported by Stanford Cancer Biology Training Program (T32 CA09302) and by a postdoctoral fellowship from the Jane Coffin Childs Memorial Fund for Medical Research. F.L.Z. was supported by a fellowship from the Agency for Science, Technology, and Research (A*STAR), Singapore. This

(C) IF for Coilin and TCAB1 in DC-patient-derived lymphoblasts.

(D) Partial PK digestion [35 S]-Met-labeled TCAB1 proteins in RRL chased with cold Met for indicated times, SDS-PAGE, top. Undigested loading, SDS-PAGE, bottom.

(E) Time course of [35 S]-Met-labeled TCAB1 proteins in RRL, chased with cold Met for indicated time, native PAGE. See (D) for loading.

(F) Co-IP of TCAB1 mutants with TERC. Western (FLAG, GAPDH) and northern (TERC, recovery) blots from co-IP of FLAG-tagged TCAB1 proteins in vitro translated in RRL and incubated with in-vitro-transcribed TERC.

(G) Co-IP of TCAB1 mutants stably expressed in HeLa S3 cells with endogenous TRiC, western blot.

(H) Co-IP of endogenous TCAB1 and endogenous TRiC from either HeLa S3 cells or DC-patient-derived lymphoblasts. Control 1, 2, and 3 are unrelated healthy donor lymphoblasts. F146L/R398 and H376Y/G435R are TCAB1 mutant patient lymphoblasts. IgG, negative control.

work was supported by NIH grants AG033747, CA125453, CA111691, and AG036695 and by the Glenn Foundation for Medical Research.

Received: June 24, 2014

Revised: September 29, 2014

Accepted: October 30, 2014

Published: November 20, 2014

REFERENCES

- Abreu, E., Aritonovska, E., Reichenbach, P., Cristofari, G., Culp, B., Terns, R.M., Lingner, J., and Terns, M.P. (2010). TIN2-tethered TPP1 recruits human telomerase to telomeres in vivo. *Mol. Cell Biol.* 30, 2971–2982.
- Albanèse, V., Yam, A.Y.-W., Baughman, J., Parnot, C., and Frydman, J. (2006). Systems analyses reveal two chaperone networks with distinct functions in eukaryotic cells. *Cell* 124, 75–88.
- Armanios, M., and Blackburn, E.H. (2012). The telomere syndromes. *Nat. Rev. Genet.* 13, 693–704.
- Artandi, S.E., and DePinho, R.A. (2010). Telomeres and telomerase in cancer. *Carcinogenesis* 31, 9–18.
- Batista, L.F., and Artandi, S.E. (2013). Understanding telomere diseases through analysis of patient-derived iPS cells. *Curr. Opin. Genet. Dev.* 23, 526–533.
- Behrends, C., Langer, C.A., Boteva, R., Böttcher, U.M., Stemp, M.J., Schaffar, G., Rao, B.V., Giese, A., Kretschmar, H., Siegers, K., and Hartl, F.U. (2006). Chaperonin TRiC promotes the assembly of polyQ expansion proteins into nontoxic oligomers. *Mol. Cell* 23, 887–897.
- Cohen, S.B., Graham, M.E., Lovrecz, G.O., Bache, N., Robinson, P.J., and Reddel, R.R. (2007). Protein composition of catalytically active human telomerase from immortal cells. *Science* 315, 1850–1853.
- Cristofari, G., and Lingner, J. (2006). Telomere length homeostasis requires that telomerase levels are limiting. *EMBO J.* 25, 565–574.
- Cristofari, G., Adolf, E., Reichenbach, P., Sikora, K., Terns, R.M., Terns, M.P., and Lingner, J. (2007). Human telomerase RNA accumulation in Cajal bodies facilitates telomerase recruitment to telomeres and telomere elongation. *Mol. Cell* 27, 882–889.
- Darzacq, X., Jádý, B.E., Verheggen, C., Kiss, A.M., Bertrand, E., and Kiss, T. (2002). Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. *EMBO J.* 21, 2746–2756.
- Diolaiti, M.E., Cimini, B.A., Kageyama, R., Charles, F.A., and Stohr, B.A. (2013). In situ visualization of telomere elongation patterns in human cells. *Nucleic Acids Res.* 41, e176.
- Egan, E.D., and Collins, K. (2012). Biogenesis of telomerase ribonucleoproteins. *RNA* 18, 1747–1759.
- Feldman, D.E., Spiess, C., Howard, D.E., and Frydman, J. (2003). Tumorigenic mutations in VHL disrupt folding in vivo by interfering with chaperonin binding. *Mol. Cell* 12, 1213–1224.
- Frydman, J., Nimmesgern, E., Erdjument-Bromage, H., Wall, J.S., Tempst, P., and Hartl, F.U. (1992). Function in protein folding of TRiC, a cytosolic ring complex containing TCP-1 and structurally related subunits. *EMBO J.* 11, 4767–4778.
- Frydman, J., Nimmesgern, E., Ohtsuka, K., and Hartl, F.U. (1994). Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones. *Nature* 370, 111–117.
- Gao, Y., Thomas, J.O., Chow, R.L., Lee, G.H., and Cowan, N.J. (1992). A cytoplasmic chaperonin that catalyzes beta-actin folding. *Cell* 69, 1043–1050.
- Hahn, W.C., Counter, C.M., Lundberg, A.S., Beijersbergen, R.L., Brooks, M.W., and Weinberg, R.A. (1999). Creation of human tumour cells with defined genetic elements. *Nature* 400, 464–468.
- Hartl, F.U., Bracher, A., and Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. *Nature* 475, 324–332.
- Holt, S.E., Aisner, D.L., Baur, J., Tesmer, V.M., Dy, M., Ouellette, M., Trager, J.B., Morin, G.B., Toft, D.O., Shay, J.W., et al. (1999). Functional requirement of p23 and Hsp90 in telomerase complexes. *Genes Dev.* 13, 817–826.
- Horn, S., Figl, A., Rachakonda, P.S., Fischer, C., Sucker, A., Gast, A., Kadel, S., Moll, I., Nagore, E., Hemminki, K., et al. (2013). TERT promoter mutations in familial and sporadic melanoma. *Science* 339, 959–961.
- Horwich, A.L., Fenton, W.A., Chapman, E., and Farr, G.W. (2007). Two families of chaperonin: physiology and mechanism. *Annu. Rev. Cell Dev. Biol.* 23, 115–145.
- Huang, F.W., Hodis, E., Xu, M.J., Kryukov, G.V., Chin, L., and Garraway, L.A. (2013). Highly recurrent TERT promoter mutations in human melanoma. *Science* 339, 957–959.
- Jádý, B.E., Bertrand, E., and Kiss, T. (2004). Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. *J. Cell Biol.* 164, 647–652.
- Killela, P.J., Reitman, Z.J., Jiao, Y., Bettegowda, C., Agrawal, N., Diaz, L.A., Jr., Friedman, A.H., Friedman, H., Gallia, G.L., Giovannella, B.C., et al. (2013). TERT promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal. *Proc. Natl. Acad. Sci. USA* 110, 6021–6026.
- Kim, Y.E., Hipp, M.S., Bracher, A., Hayer-Hartl, M., and Hartl, F.U. (2013). Molecular chaperone functions in protein folding and proteostasis. *Annu. Rev. Biochem.* 82, 323–355.
- Kitamura, A., Kubota, H., Pack, C.-G., Matsumoto, G., Hirayama, S., Takahashi, Y., Kimura, H., Kinjo, M., Morimoto, R.I., and Nagata, K. (2006). Cytosolic chaperonin prevents polyglutamine toxicity with altering the aggregation state. *Nat. Cell Biol.* 8, 1163–1170.
- Kubota, H., Hynes, G.M., Kerr, S.M., and Willison, K.R. (1997). Tissue-specific subunit of the mouse cytosolic chaperonin-containing TCP-1. *FEBS Lett.* 402, 53–56.
- Liu, T., Daniels, C.K., and Cao, S. (2012). Comprehensive review on the HSC70 functions, interactions with related molecules and involvement in clinical diseases and therapeutic potential. *Pharmacol. Ther.* 136, 354–374.
- Mitchell, J.R., Wood, E., and Collins, K. (1999). A telomerase component is defective in the human disease dyskeratosis congenita. *Nature* 402, 551–555.
- Morimoto, R.I. (2008). Proteotoxic stress and inducible chaperone networks in neurodegenerative disease and aging. *Genes Dev.* 22, 1427–1438.
- Nandakumar, J., Bell, C.F., Weidenfeld, I., Zaug, A.J., Leinwand, L.A., and Cech, T.R. (2012). The TEL patch of telomere protein TPP1 mediates telomerase recruitment and processivity. *Nature* 492, 285–289.
- Palm, W., and de Lange, T. (2008). How shelterin protects mammalian telomeres. *Annu. Rev. Genet.* 42, 301–334.
- Pfeiffer, V., and Lingner, J. (2013). Replication of telomeres and the regulation of telomerase. *Cold Spring Harb. Perspect. Biol.* 5, a010405.
- Rashid, R., Liang, B., Baker, D.L., Youssef, O.A., He, Y., Phipps, K., Terns, R.M., Terns, M.P., and Li, H. (2006). Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita. *Mol. Cell* 21, 249–260.
- Sexton, A.N., Youmans, D.T., and Collins, K. (2012). Specificity requirements for human telomere protein interaction with telomerase holoenzyme. *J. Biol. Chem.* 287, 34455–34464.
- Stern, J.L., Zyner, K.G., Pickett, H.A., Cohen, S.B., and Bryan, T.M. (2012). Telomerase recruitment requires both TCAB1 and Cajal bodies independently. *Mol. Cell Biol.* 32, 2384–2395.
- Tam, S., Geller, R., Spiess, C., and Frydman, J. (2006). The chaperonin TRiC controls polyglutamine aggregation and toxicity through subunit-specific interactions. *Nat. Cell Biol.* 8, 1155–1162.
- Taylor, R.C., and Dillin, A. (2011). Aging as an event of proteostasis collapse. *Cold Spring Harb. Perspect. Biol.* 3, a004440.
- Tycowski, K.T., Shu, M.D., Kukoyi, A., and Steitz, J.A. (2009). A conserved WD40 protein binds the Cajal body localization signal of scaRNP particles. *Mol. Cell* 34, 47–57.

- Vainberg, I.E., Lewis, S.A., Rommelaere, H., Ampe, C., Vandekerckhove, J., Klein, H.L., and Cowan, N.J. (1998). Prefoldin, a chaperone that delivers unfolded proteins to cytosolic chaperonin. *Cell* 93, 863–873.
- Venteicher, A.S., Meng, Z., Mason, P.J., Veenstra, T.D., and Artandi, S.E. (2008). Identification of ATPases pontin and reptin as telomerase components essential for holoenzyme assembly. *Cell* 132, 945–957.
- Venteicher, A.S., Abreu, E.B., Meng, Z., McCann, K.E., Terns, R.M., Veenstra, T.D., Terns, M.P., and Artandi, S.E. (2009). A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. *Science* 323, 644–648.
- Wolff, S., Weissman, J.S., and Dillin, A. (2014). Differential scales of protein quality control. *Cell* 157, 52–64.
- Wong, J.M.Y., and Collins, K. (2006). Telomerase RNA level limits telomere maintenance in X-linked dyskeratosis congenita. *Genes Dev.* 20, 2848–2858.
- Yam, A.Y., Xia, Y., Lin, H.-T.J., Burlingame, A., Gerstein, M., and Frydman, J. (2008). Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies. *Nat. Struct. Mol. Biol.* 15, 1255–1262.
- Zaug, A.J., Crary, S.M., Jesse Fioravanti, M., Campbell, K., and Cech, T.R. (2013). Many disease-associated variants of hTERT retain high telomerase enzymatic activity. *Nucleic Acids Res.* 41, 8969–8978.
- Zhang, J.H., Chung, T.D., and Oldenburg, K.R. (1999). A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J. Biomol. Screen.* 4, 67–73.
- Zhong, F., Savage, S.A., Shkreli, M., Giri, N., Jessop, L., Myers, T., Chen, R., Alter, B.P., and Artandi, S.E. (2011). Disruption of telomerase trafficking by TCAB1 mutation causes dyskeratosis congenita. *Genes Dev.* 25, 11–16.
- Zhong, F.L., Batista, L.F.Z., Freund, A., Pech, M.F., Venteicher, A.S., and Artandi, S.E. (2012). TPP1 OB-fold domain controls telomere maintenance by recruiting telomerase to chromosome ends. *Cell* 150, 481–494.
- Zhu, Y., Tomlinson, R.L., Lukowiak, A.A., Terns, R.M., and Terns, M.P. (2004). Telomerase RNA accumulates in Cajal bodies in human cancer cells. *Mol. Biol. Cell* 15, 81–90.

Leptin Mediates the Increase in Blood Pressure Associated with Obesity

Stephanie E. Simonds,¹ Jack T. Pryor,^{2,3} Eric Ravussin,⁴ Frank L. Greenway,⁴ Ralph Dileone,⁵ Andrew M. Allen,⁶ Jaspreet Bassi,⁶ Joel K. Elmquist,⁷ Julia M. Keogh,⁸ Elana Henning,⁸ Martin G. Myers, Jr.,⁹ Julio Licinio,¹⁰ Russell D. Brown,¹ Pablo J. Enriori,¹ Stephen O'Rahilly,⁸ Scott M. Sternson,¹¹ Kevin L. Grove,¹² David C. Spanswick,^{1,2} I. Sadaf Farooqi,^{8,*} and Michael A. Cowley^{1,*}

¹Department of Physiology, Monash Obesity and Diabetes Institute, Monash University, Clayton, VIC 3800, Australia

²Warwick Medical School, University of Warwick, Coventry CV4 7AL, UK

³Neurosolutions Ltd., Coventry CV4 7ZS, UK

⁴The Pennington Biomedical Research Center, Baton Rouge, LA 70808, USA

⁵Department of Psychiatry, Yale School of Medicine, New Haven, CT 06511, USA

⁶Department of Physiology, The University of Melbourne, Victoria 3010, Australia

⁷Division of Endocrinology & Metabolism, Department of Internal Medicine, The University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

⁸University of Cambridge Metabolic Research Laboratories, Wellcome Trust-MRC Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK

⁹Division of Metabolism, Endocrinology, and Diabetes, Department of Internal Medicine, University of Michigan, Ann Arbor, MI 48109, USA

¹⁰Mind and Brain Theme, South Australian Health and Medical Research Institute and Department of Psychiatry, School of Medicine, Flinders University, Adelaide, SA 5001, Australia

¹¹Janelia Farm Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA

¹²Division of Diabetes, Obesity & Metabolism, Oregon National Primate Research Center, Portland, OR 97239, USA

*Correspondence: isf20@cam.ac.uk (I.S.F.), michael.cowley@monash.edu (M.A.C.)

<http://dx.doi.org/10.1016/j.cell.2014.10.058>

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

SUMMARY

Obesity is associated with increased blood pressure (BP), which in turn increases the risk of cardiovascular diseases. We found that the increase in leptin levels seen in diet-induced obesity (DIO) drives an increase in BP in rodents, an effect that was not seen in animals deficient in leptin or leptin receptors (LepR). Furthermore, humans with loss-of-function mutations in leptin and the LepR have low BP despite severe obesity. Leptin's effects on BP are mediated by neuronal circuits in the dorsomedial hypothalamus (DMH), as blocking leptin with a specific antibody, antagonist, or inhibition of the activity of LepR-expressing neurons in the DMH caused a rapid reduction of BP in DIO mice, independent of changes in weight. Re-expression of LepRs in the DMH of DIO LepR-deficient mice caused an increase in BP. These studies demonstrate that leptin couples changes in weight to changes in BP in mammalian species.

INTRODUCTION

Obesity increases the risk for hypertension and is a major driver of morbidity and mortality due to cardiovascular diseases (Du-

stan, 1983; Poirier et al., 2006). Studies in rodents with diet-induced obesity (DIO) suggest that increased sympathetic nerve activity (SNA) is an important mediator of obesity-induced hypertension as α and β adrenergic receptor antagonists and renal denervation significantly blunt the rise in blood pressure (BP) associated with weight gain (Carlyle et al., 2002; Esler et al., 2006; Kassab et al., 1995). However, the precise molecular and neural mechanisms that link changes in weight with changes in BP have not been fully elucidated.

Circulating concentrations of the adipocyte-derived hormone leptin increase in proportion to adipose tissue mass and fall with weight loss (Considine et al., 1996; Maffei et al., 1995). As such, we hypothesized that leptin may be involved in coupling changes in body weight (BW) to changes in BP. Leptin regulates energy homeostasis by acting on hypothalamic neuronal circuits expressing the signaling isoform of the leptin receptor (LepR) to reduce calorie intake and increase energy expenditure (Halaas et al., 1997; Harris et al., 1998; Maffei et al., 1995; Zhang et al., 1994). Leptin can increase SNA, leading to increases in BP and heart rate (HR) (Haynes, 2000; Mark et al., 1999). In the arcuate nucleus of the hypothalamus (ARH), leptin stimulates the expression of pro-opiomelanocortin (POMC) and increases the activity of POMC neurons, which release the POMC peptides (α , β , and γ melanocyte-stimulating hormones [MSHs]) that act on melanocortin 4-receptor (MC4R)-expressing neurons in the paraventricular nucleus of the hypothalamus (PVH) and other brain regions to increase SNA (Cone, 2005; Cowley et al., 1999, 2001; Haynes et al., 1999). However,

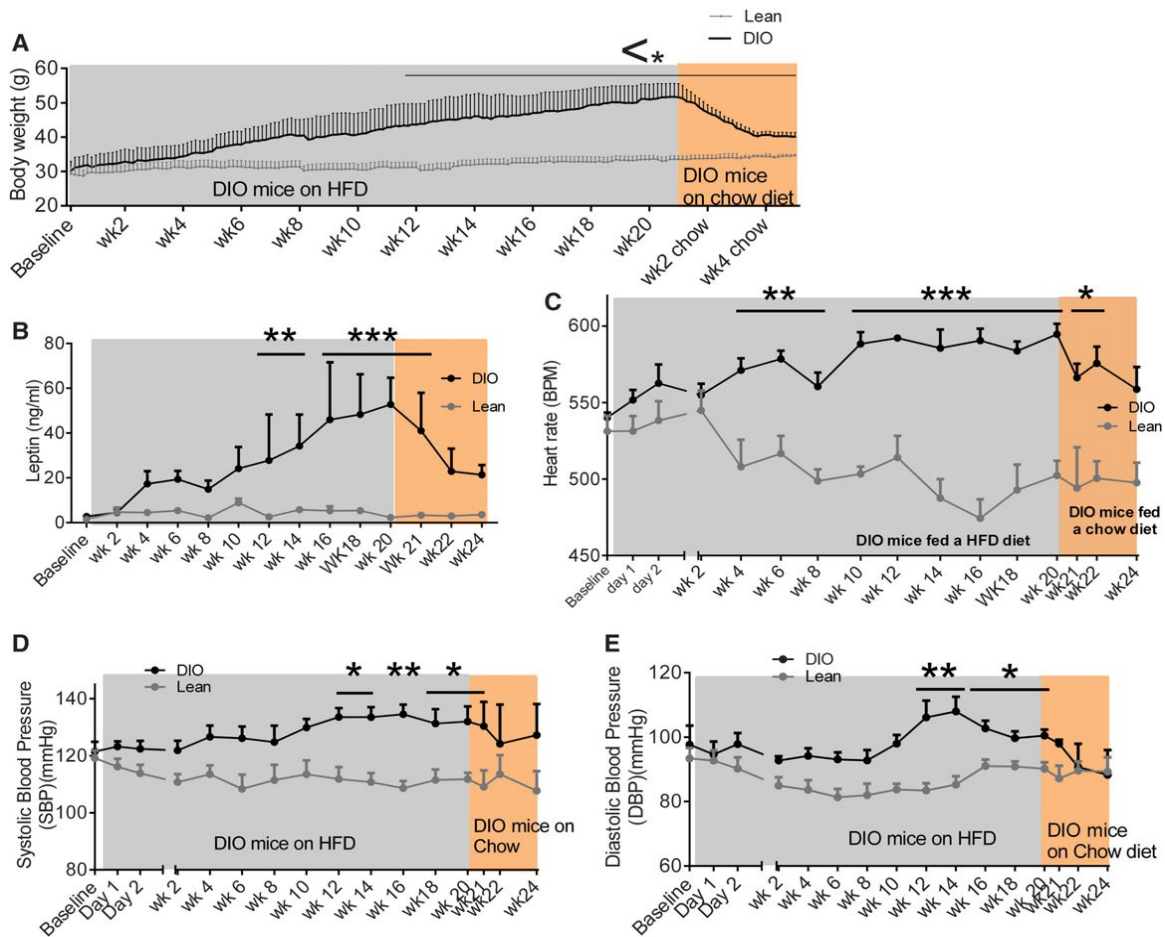


Figure 1. Temporal Relationship between Changes in Weight Gain, Plasma Leptin Concentrations, and Blood Pressure

(A) Body weight (g), of mice from 4 weeks of age ad libitum fed either a high-calorie diet (DIO mice) or chow diet (lean mice) for 20 weeks. Following this 20 week period, DIO mice were switched to chow diet for 4 weeks; $n = 3-7$.

(B-E) (B) Plasma leptin concentration (ng/ml), (C) heart rate BPM (beats per min), (D) SBP, and (E) DBP (mmHg) at 2 week intervals, beginning at baseline (4-week-old mice). Following this baseline period, mice were fed a high-calorie diet (DIO) or chow diet (lean) for 20 weeks ad libitum. After 20 weeks of HFD in DIO mice, these mice were switched to chow diet for 4 weeks. $n = 3-18$.

Values mean \pm SEM; two-way AVOVA, Bonferroni post hoc test. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

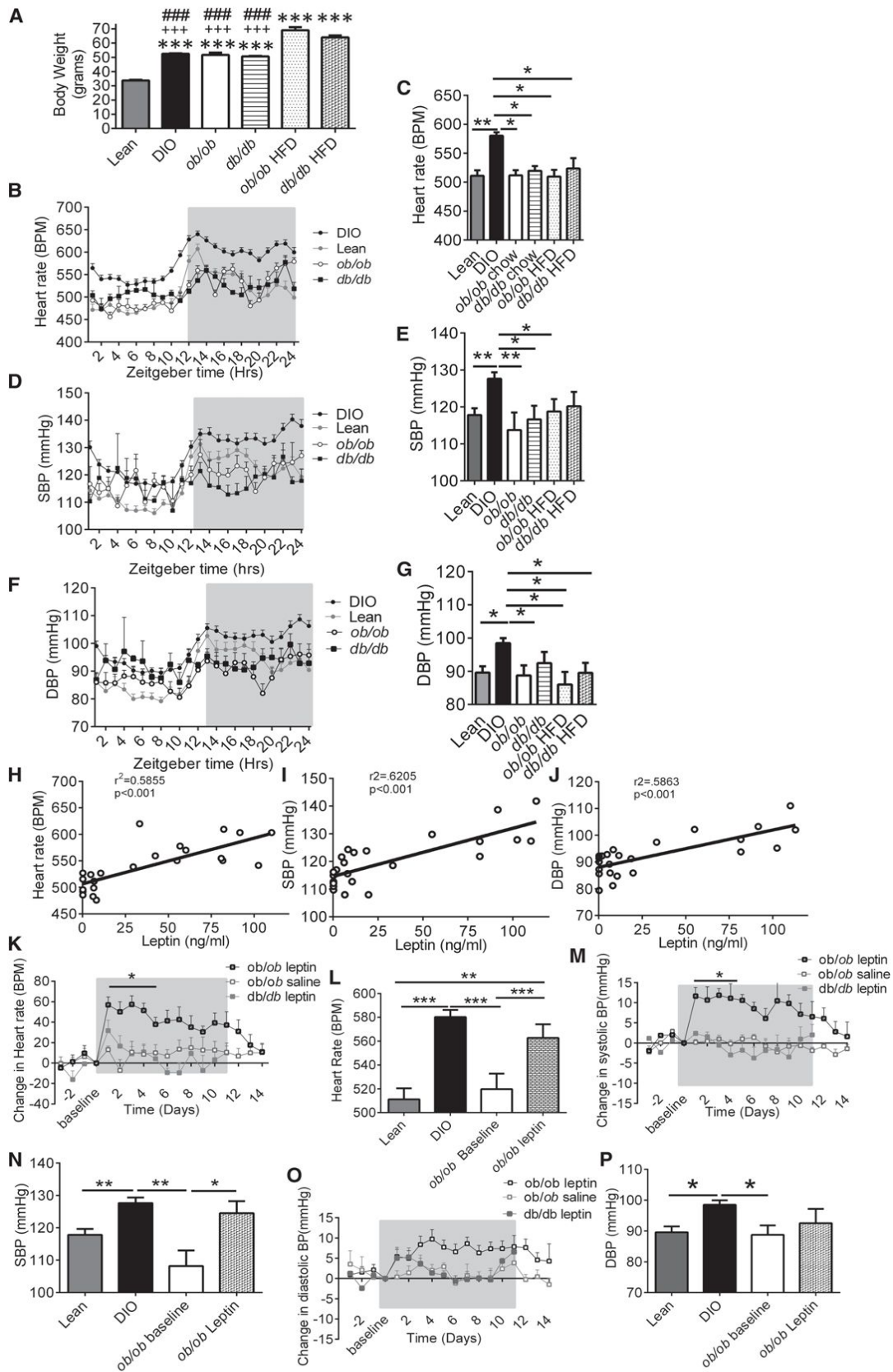
POMC neurons become unresponsive to leptin in obesity, and leptin can act independently of MC4R signaling (Enriori et al., 2011; Patterson et al., 2011; Scott et al., 2009). Therefore, leptin's effects beyond the melanocortin circuits need to be investigated. The dorsomedial hypothalamus (DMH) is critical for leptin's ability to regulate brown adipose tissue (BAT) temperature and the cardiovascular system (Enriori et al., 2011; Fontes et al., 2001; Horiuchi et al., 2006; Marsh et al., 2003; Rezai-Zadeh et al., 2014).

Here, we investigated the development of obesity-induced hypertension. We demonstrate that, in DIO mice, increasing levels of leptin directly lead to an increase in HR and BP and that blocking the actions of leptin reverses these effects via neural circuits originating in the DMH. Furthermore, humans with loss-of-function mutations in leptin and its receptor have normal BP despite severe obesity, suggesting that these mechanisms are likely to be preserved in humans.

RESULTS

Weight Gain Increases Leptin Levels, Heart Rate, and Blood Pressure

The temporal association between weight gain, changes in circulating leptin levels, HR, and BP were first examined. Four-week-old C57Bl/6J mice on a chow diet were implanted with radiotelemetric BP probes; baseline measurements were recorded at 6 weeks. As BW increased (Figure 1A), plasma leptin levels also progressively increased (Figure 1B). After 4 weeks of HFD, HR became significantly elevated (Figure 1C) and remained elevated throughout the 20 week recording period compared to chow fed mice. Systolic blood pressure (SBP) and diastolic blood pressure (DBP) were significantly elevated after 12 weeks of HFD (Figures 1D and 1E), subsequent to the rise in plasma leptin concentration. When HFD-fed mice were returned to a chow diet following 20 weeks of HFD feeding, mice



(legend on next page)

lost 3.9 g in the first week, 3.9 g in the second week, 3.4 g in the third week, and 0.4 g in the fourth week (Figure 1A). DBP reduced after 1 week (Figure 1E), and SBP (Figure 1D) and HR (Figure 1C) reduced after 2 weeks of chow feeding. These findings suggest that leptin appears to increase before HR and BP increases in DIO, effects that are reversed with weight loss.

Leptin Signaling Is Required for Obesity-Induced Increases in Blood Pressure

We examined whether the effects seen in DIO mice are dependent upon leptin signaling, using mice lacking leptin (*ob/ob*) or the LepR (*db/db*). HFD-fed *ob/ob* and *db/db* mice were significantly ($p < 0.001$) heavier than HFD-fed DIO and chow-fed lean, *ob/ob*, and *db/db* mice (Figure 2A). Despite the increased severity of obesity in the *ob/ob* and *db/db* mice, only DIO mice exhibited elevated HR, SBP, and DBP (Figures 2B–2G), demonstrating that increased BW alone is not the cause of the increase in BP in obesity. There was a greater difference in HR and BP between DIO mice to other mice during the dark period compared to the light period (Figures S1A–S1C available online). The increase in cardiovascular parameters in DIO mice was not due to increased activity (Figures S1D and S1E).

Strong correlations were found in C57Bl/6J mice between plasma leptin concentration and HR (Figure 2H), SBP (Figure 2I), and DBP (Figure 2J). No correlation was found between plasma insulin concentration and HR, SBP, or DBP (Figures S1G–S1I). To examine the direct effect of leptin on BP, leptin-deficient *ob/ob* mice were treated for 11 days with either leptin or vehicle. After 11 days of treatment, the average plasma leptin concentra-

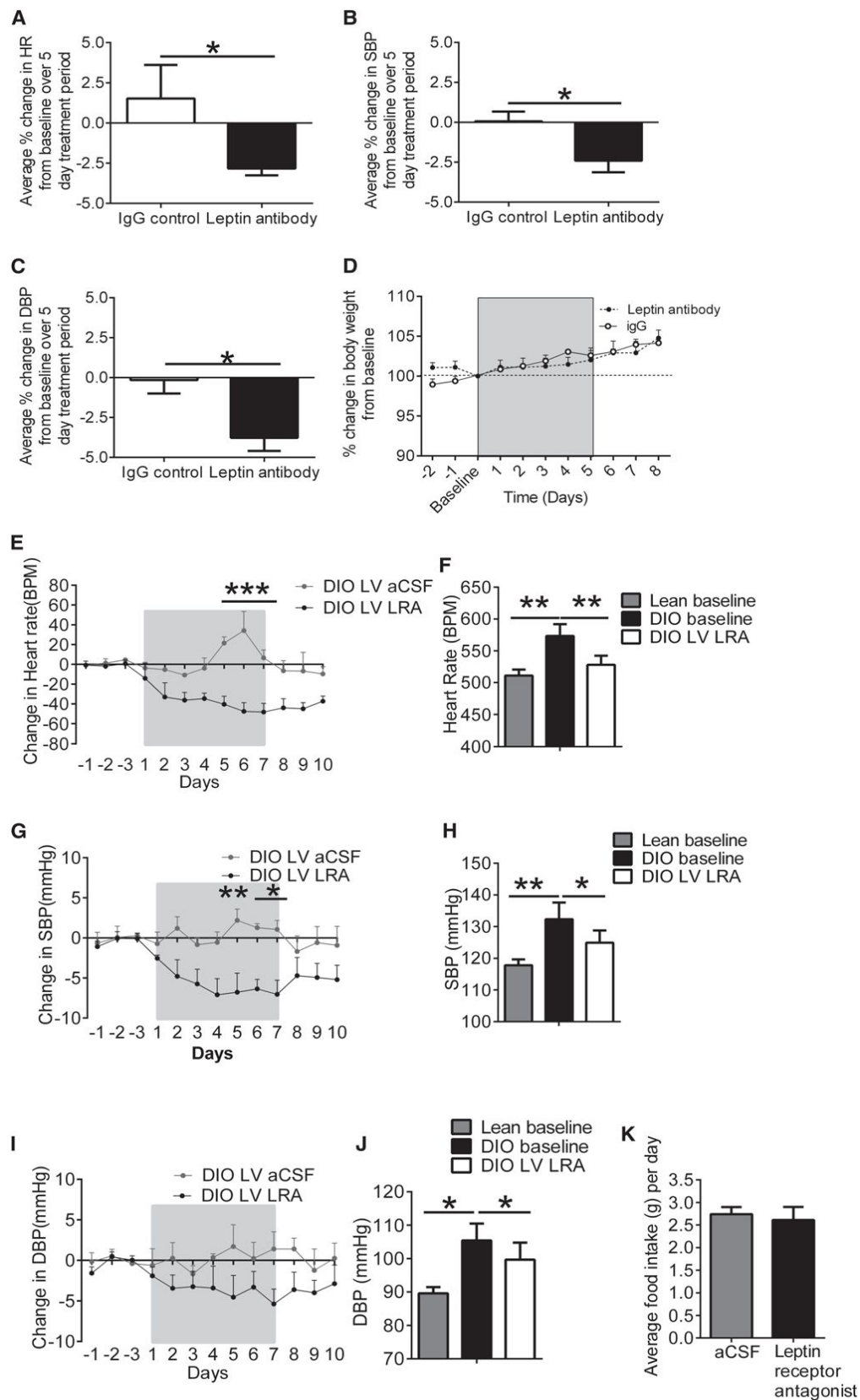
tion of *ob/ob* mice treated with leptin was 27.7 ± 2.9 ng/ml compared to 36.3 ± 4.9 ng/ml in DIO mice and 6.6 ± 1.1 ng/ml in lean mice (Figure S1F). Exogenous leptin treatment in *ob/ob* mice progressively decreased BW (Figure S1M) and food intake (FI) (Figure S1N) compared to leptin-treated *db/db* mice and vehicle-treated *ob/ob* mice. Despite this decrease in BW, leptin treatment increased HR (Figures 2K and 2L) and SBP (Figures 2M and 2N) in leptin-treated *ob/ob* mice compared to leptin-treated *db/db* mice and vehicle-treated *ob/ob* mice. DBP also increased in *ob/ob* leptin-treated mice, however, not to the same extent as SBP (Figures 2O and 2P). The magnitude of the increase in HR and BP in leptin-treated *ob/ob* mice decreased with time, coincident with their substantial reduction in BW. The 24 hr circadian rhythm of HR (Figure S2J), SBP (Figure S2K), and DBP (Figure S1L) of leptin-treated *ob/ob* mice was increased toward the baseline recordings of DIO mice. These changes were not explained by increased activity (Figure S1O).

Inhibition of Leptin Signaling Reduces Elevated BP

Peripheral administration of a leptin antibody into DIO mice for 5 days caused a significant reduction in HR (Figure 3A), SBP (Figure 3B), and DBP (Figure 3C), although no change in BW was observed (Figure 3D). Central administration of a leptin receptor antagonist (LRA) into the lateral ventricle (LV) of hypertensive DIO mice over a 7 day period significantly reduced the elevated HR ($p < 0.001$; Figure 3E). By the seventh day of LRA treatment, HR was comparable to the average HR of lean animals (Figures 3F and S2A). Systolic BP also progressively decreased when DIO mice were treated with the LRA, and by day 5 of LRA

Figure 2. Leptin Signaling Is Required for the Obesity-Induced Increases in BP

(A) Body weight (g) of mice fed 20 weeks a high-calorie (DIO, *ob/ob*, and *db/db* mice) or chow (lean *ob/ob* and *db/db* mice) diet ad libitum, *** $p < 0.001$ versus lean, *** $p < 0.001$ versus *ob/ob* HFD fed, ### $p < 0.001$ versus *db/db* HFD (lean, $n = 14$; DIO, $n = 37$; *ob/ob*, $n = 8$; *db/db*, $n = 3$; *ob/ob*, HFD = 3; *db/db*, HFD = 5).
 (B) Average heart rate (BPM) over a 24 hr period of mice fed a high-calorie (DIO mice) or chow (lean mice) diet, *ob/ob* mice and *db/db* mice fed a chow diet for over 20 weeks (lean, $n = 14$; DIO, $n = 37$; *ob/ob*, $n = 8$; *db/db*, $n = 3$; *ob/ob*, HFD = 3; *db/db*, HFD = 5).
 (C) Heart rate (BPM) over a 24 hr period in mice fed a chow diet (lean mice), HFD-fed DIO mice, chow-fed *ob/ob* and *db/db* mice, and HFD-fed *ob/ob* and *db/db* mice. Measurements at 20 weeks are presented (lean, $n = 14$; DIO, $n = 37$; *ob/ob*, $n = 8$; *db/db*, $n = 3$; *ob/ob*, HFD = 3; *db/db*, HFD = 5).
 (D) Average SBP (mmHg) over a 24 hr period of mice fed a high-calorie (DIO mice) or chow (lean mice) diet, *ob/ob* mice and *db/db* mice fed a chow diet for 20 weeks (lean, $n = 14$; DIO, $n = 37$; *ob/ob*, $n = 8$; *db/db*, $n = 3$; *ob/ob*, HFD = 3; *db/db*, HFD = 5).
 (E) SBP (mmHg) over a 24 hr period in mice fed a chow diet (lean mice), HFD-fed DIO mice, chow-fed *ob/ob* and *db/db* mice, and HFD-fed *ob/ob* and *db/db* mice for 20 weeks (lean, $n = 14$; DIO, $n = 37$; *ob/ob*, $n = 8$; *db/db*, $n = 3$; *ob/ob*, HFD = 3; *db/db*, HFD = 5).
 (F) Average DBP (mmHg) over a 24 hr period of mice fed a high-calorie (DIO mice) or chow (lean mice) diet, *ob/ob* mice and *db/db* mice fed a chow diet for 20 weeks (lean, $n = 14$; DIO, $n = 37$; *ob/ob*, $n = 8$; *db/db*, $n = 3$; *ob/ob*, HFD = 3; *db/db*, HFD = 5).
 (G) DBP (mmHg) over a 24 hr period in mice fed a chow diet (lean mice), HFD-fed DIO mice, chow-fed *ob/ob* and *db/db* mice, and HFD-fed *ob/ob* and *db/db* mice (lean, $n = 14$; DIO, $n = 37$; *ob/ob*, $n = 8$; *db/db*, $n = 3$; *ob/ob*, HFD = 3; *db/db*, HFD = 5).
 (H) Plasma leptin (ng/ml) correlates with heart rate (BPM). Linear regression, $r^2 = 0.5855$; $p < 0.001$; $n = 24$.
 (I) Plasma leptin (ng/ml) correlates with SBP (mmHg). Linear regression, $r^2 = 0.6205$; $p < 0.001$; $n = 24$.
 (J) Plasma leptin (ng/ml) correlates with DBP. Linear regression, $r^2 = 0.5883$; $p < 0.001$; $n = 24$.
 (K) Exogenous leptin administration into *ob/ob* mice increased heart rate. Vehicle (Veh) treatment into *ob/ob* mice and leptin administration into *db/db* mice produced no change. Two-way ANOVA, Bonferroni post hoc test. $n = 3$ –8.
 (L) Average heart rate (BPM) over 11 day treatment period of leptin-treated *ob/ob* mice, compared to the baseline HR of *ob/ob* mice, the baseline of DIO and lean mice. One-way ANOVA, Bonferroni post hoc test, paired t test between leptin-treated *ob/ob* mice and *ob/ob* baseline, $n = 8$ –37.
 (M) Exogenous leptin administration into *ob/ob* mice, increased SBP (mmHg). Vehicle treatment into *ob/ob* mice and leptin administration into *db/db* mice produced no change. Values represent mean \pm SEM. Two-way ANOVA, Bonferroni post hoc test. $n = 3$ –8.
 (N) Average SBP (mmHg) over 11 day treatment period of leptin-treated *ob/ob* mice compared to *ob/ob* baseline SBP, DIO, and lean baseline SBP measurements. One-way ANOVA, Bonferroni post hoc test, paired t test between leptin-treated *ob/ob* mice and *ob/ob* baseline. $n = 8$ –37.
 (O) Exogenous leptin administration into *ob/ob* mice increased diastolic BP (DBP) (mmHg). Vehicle treatment into *ob/ob* mice and leptin administration into *db/db* mice produced no change. Two-way ANOVA, Bonferroni post hoc test. $n = 3$ –8.
 (P) Average DBP (mmHg) over 11 day treatment period of leptin-treated *ob/ob* mice compared to *ob/ob* baseline DBP, DIO baseline DBP, and lean mice DBP. One-way ANOVA, paired t test between leptin-treated *ob/ob* mice and *ob/ob* baseline. $n = 8$ –37.
 Values represent mean \pm SEM, one-way ANOVA, Bonferroni post hoc test, unless specifically stated. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. See also Figure S1.



(legend on next page)

treatment, SBP was significantly ($p \leq 0.05$) decreased compared to vehicle-treated DIO mice (Figure 3G). By the seventh day of LRA treatment, SBP was significantly ($p < 0.05$) lower compared to the SBP at baseline in DIO mice and was comparable to the baseline SBP of lean mice (Figures 3H and S2B). SBP increased after LRA treatment ended. DBP also decreased throughout the LRA treatment period (Figures 3I, 3J, and S2C). These changes were seen despite the absence of a significant change in FI (Figure 3K), indicating that these changes are mediated by blocking the effects of leptin signaling rather than through changes in FI and BW.

Leptin Modulates BP via Neural Circuits in the Dorsomedial Hypothalamus

To determine whether leptin's effects on BP and HR were mediated by neurons in the DMH, the LRA was injected directly into the DMH of hypertensive DIO mice daily over 7 days. LRA treatment in the DMH decreased HR (Figures 4A, 4B, and S3A) and SBP (Figures 4C, 4D, and S3B) by day 7 of treatment. No significant change in DBP (Figures 4E, 4F, and S3C) was observed. Furthermore, injection of an AAV expressing a short hairpin RNA directed against the LepR (Hommel et al., 2006) in the DMH of DIO mice led to a sustained decrease in SBP (Figures 4G, 4H, and S3D) after 4 weeks. Utilizing mice in which the LepR is flanked by loxP sites and can be deleted by an AAV Cre, HFD for 20 weeks induced increased SBP at baseline, whereas AAV Cre administration into the DMH (hence deleting the LepR in only the DMH regions) decreased SBP (Figure 4I). Additionally, we examined the effects of reactivation of the LepR by injection of an AAV-expressing Cre recombinase into the DMH of normotensive morbidly obese LepR null mice (expressing a loxP flanked STOP codon in front of the LepR, called LepR transcriptional blocker or "LepR TB" mice; Berglund et al., 2012). Both HR (Figure 4J) and SBP (Figure 4K) increased rapidly after LepR expression was reactivated in just the DMH of obese LepR-deficient mice.

Depolarization of DMH LepR-Expressing Neurons

To determine the electrophysiological effect of leptin on DMH neurons, we recorded the electrical activity of neurons expressing LepR in LepR-Cre-YFP mice (Leshan et al., 2012). Whole-cell recordings were obtained from 34 DMH neurons expressing LepR. Application of 100 nM leptin induced membrane

depolarization and/or an increase in spontaneous action potential firing rate in 38.2% of recorded neurons (Figure 5A). Leptin induced a mean peak amplitude depolarization of 4.9 ± 1.0 mV from a mean resting potential of -49.9 ± 3.1 mV to a new steady-state membrane potential of -45.0 ± 2.7 mV ($p = 0.0004$). Leptin-induced excitation was associated with an increase in firing rate from a mean control basal level of 0.48 ± 0.28 Hz to 0.64 ± 0.34 Hz in the presence of leptin ($n = 5$), effects that were at least in part reversible after washout of leptin ($p = 0.59$). In two neurons, leptin induced sub- and supra-threshold oscillations in membrane potential (Figure S4A). Voltage-current relations, generated in response to a range of depolarizing and hyperpolarizing rectangular-wave current pulses (-150 to $+100$ pA, 1,000 ms duration) revealed that leptin-induced excitation was principally associated with a trend for a decrease ($p = 0.86$) in neuronal input resistance from 841 ± 111 M Ω in the absence to 743 ± 138 M Ω in the presence of leptin. Plots of the voltage-current relations revealed extrapolated reversal potentials for leptin-induced excitation around -35 mV (Figure 5B). Taken together, these data suggest that leptin-induced excitation is mediated via activation of one or more nonselective cation conductances. In addition to these effects on membrane input conductance, leptin also induced modulation of intrinsic subthreshold active conductance in a subpopulation of DMH neurons. In DMH neurons, membrane depolarization from negative holding potentials (< -65 mV) or depolarization at the offset of the membrane response to hyperpolarization from potentials close to resting potential (-45 to -50 mV) evoked a transient depolarizing potential consistent with activation of a low-threshold T-type calcium conductance. In the presence of leptin, this potential was prolonged (Figure 5B), the half-time to decay increasing from 132 ± 59 ms in the absence to 179 ± 61 ms in the presence of leptin. These data are consistent with leptin modulating intrinsic active conductances in a subpopulation of DMH neurons. In addition to these postsynaptic effects, leptin induced an increase in spontaneous excitatory postsynaptic potentials (EPSPs) in a subpopulation ($n = 2$) of DMH neurons (Figure S4B). The mean frequency of spontaneous EPSPs increased from 0.07 ± 0.02 Hz to 0.31 ± 0.11 Hz in the presence of leptin. These EPSPs could be of sufficient magnitude to reach threshold for firing, suggesting that indirect presynaptic effects of leptin on DMH neurons can contribute to leptin-induced increases in neuronal excitability.

Figure 3. Blockade of Leptin Actions in DIO Mice Reduces Elevated BP and HR

(A–C) (A) Percentage change in HR, (B) SBP, and (C) DBP over a 5 day treatment period with either leptin antibody or vehicle (IgG control). Student's *t* test, $n = 4$ –8. (D) Percentage daily change in body weight of DIO mice before and during 5 days of treatment and posttreatment of leptin antibody or vehicle (IgG control). Two-way ANOVA, Bonferroni post hoc test. $n = 4$ –8. (E) Daily change in HR (BPM) in DIO mice treated LV with aCSF or leptin receptor antagonist (LRA). Two way ANOVA, Bonferroni post hoc, $n = 4$ –6. (F) Total 24 hr HR(BPM) after 7 days of LRA treatment in DIO mice compared to lean and DIO mice baselines. *t* test paired (DIO baseline versus LRA), unpaired *t* test. $n = 6$ –37. (G) Daily change in SBP (mmHg) from baseline in DIO mice treated in the LV with aCSF or LRA. Two-way ANOVA, Bonferroni post hoc test, $n = 4$ –6. (H) Total 24 hr SBP (mmHg) after 7 days of LRA treatment in DIO mice compared to the baselines of lean and DIO mice. *t* test paired (DIO baseline versus LRA), unpaired *t* test. $n = 6$ –37. (I) Daily change in DBP (mmHg) in DIO mice treated with LVaCSF or LRA. Two-way ANOVA, Bonferroni post hoc, $n = 4$ –6. (J) Total 24 hr DBP (mmHg) after 7 days of LRA in DIO mice compared to the baseline measurements of lean and DIO mice. *t* test paired (DIO baseline versus LRA), unpaired *t* test. $n = 6$ –37. (K) Average food intake over the 7 day treatment period of LVaCSF (control) or LRA DIO mice. *t* test, $n = 4$ –6. Values represent mean \pm SEM. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. See also Figure S2.

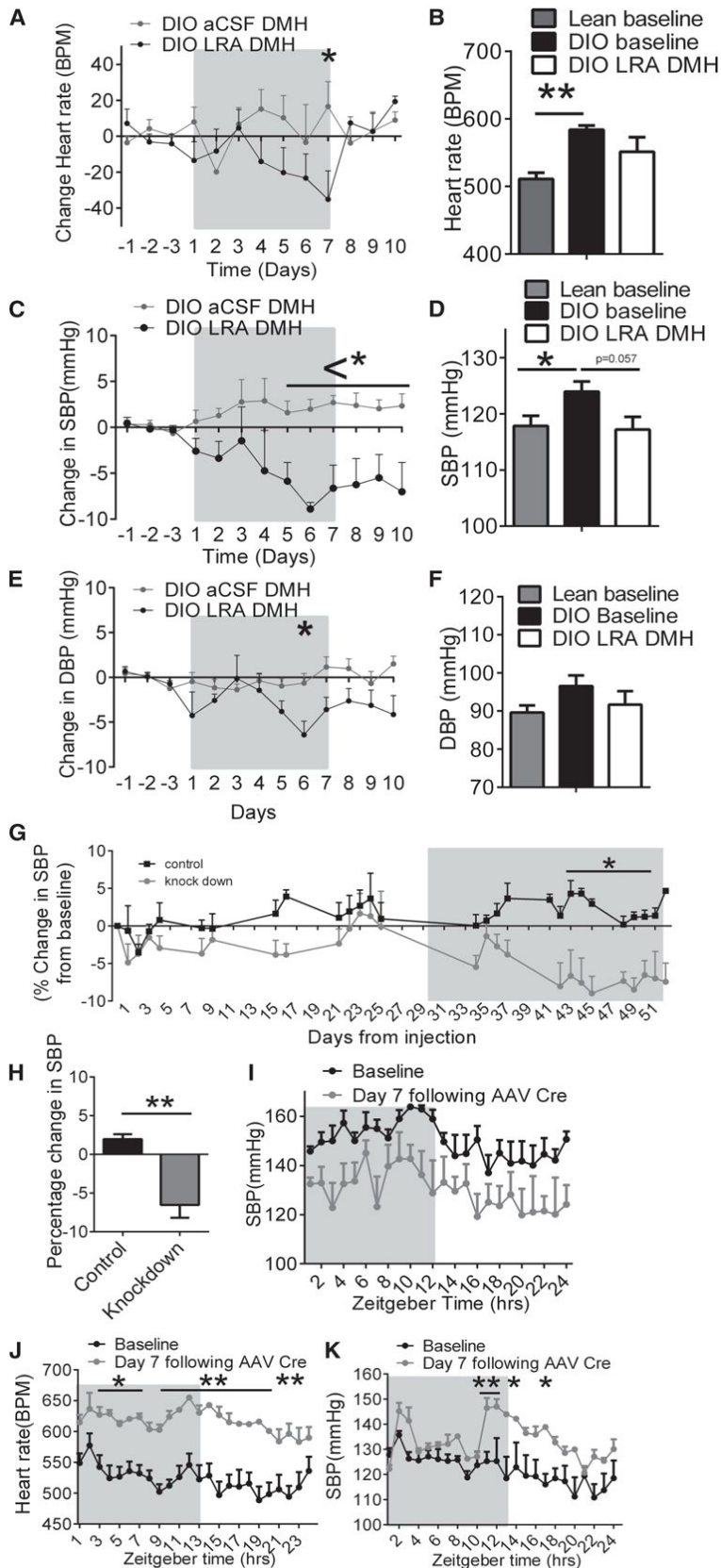


Figure 4. The DMH Is Involved in Leptin-Mediated Increases in HR and BP

(A) Daily change in HR (BPM) from baseline in DIO mice treated in the DMH with aCSF or a LRA. Two-way ANOVA, Bonferroni post hoc test, $n = 4$.

(B) Total 24 hr HR (BPM) after 7 days of LRA treatment in the DMH of DIO mice compared to lean and DIO mice at baseline. t test paired (DIO baseline versus LRA), unpaired t test. $n = 4-37$.

(C) Daily change in SBP (mmHg) from baseline in DIO mice treated in the DMH with aCSF or a LRA. Two-way ANOVA, Bonferroni post hoc test, $n = 4$.

(D) Total 24 hr SBP (mmHg) after 7 days of LRA treatment in the DMH of DIO mice compared to lean and DIO mice at baseline. t test paired (DIO baseline versus LRA), unpaired t test. $n = 4-37$.

(E) Daily change in DBP (mmHg) from baseline in DIO mice treated in the DMH with aCSF or a LRA. Two-way ANOVA, Bonferroni post hoc test, $n = 4$.

(F) 24 hr DBP (mmHg) after 7 days of LRA treatment in the DMH of DIO mice compared to lean and DIO mice at baseline. t test paired (DIO baseline versus LRA), unpaired t test. $n = 4-37$.

(G) Percentage change in SBP (mmHg) of DIO mice, following control (scrambled AAV) or leptin receptor knockdown AAV into the DMH. Two-way ANOVA, Bonferroni post hoc test, $n = 4-6$.

(H) Quantitative analysis from day 31 onward of AAV scrambled control-injected DIO mice and AAV LepR knockdown-treated DIO mice. t test, $n = 4-6$.

(I) 24 hr circadian rhythm of SBP (mmHg) at baseline and then following the seventh day of AAV Cre injection into the DMH of LepR flox mice. Two-way ANOVA, Bonferroni post hoc test, $n = 4$.

(J) 24 hr circadian rhythm of heart rate (BPM) at baseline and then following the seventh day of AAV Cre injection into the DMH of LepR flox TB mice. Two-way ANOVA, Bonferroni post hoc test, $n = 4$.

(K) 24 hr circadian rhythm of SBP (mmHg) at baseline and then following the seventh day of AAV Cre injection into the DMH of LepR flox TB mice. Two-way ANOVA, Bonferroni post hoc test, $n = 4$.

Values represent mean \pm SEM. * $p < 0.05$ and ** $p < 0.01$. See also Figure S3.

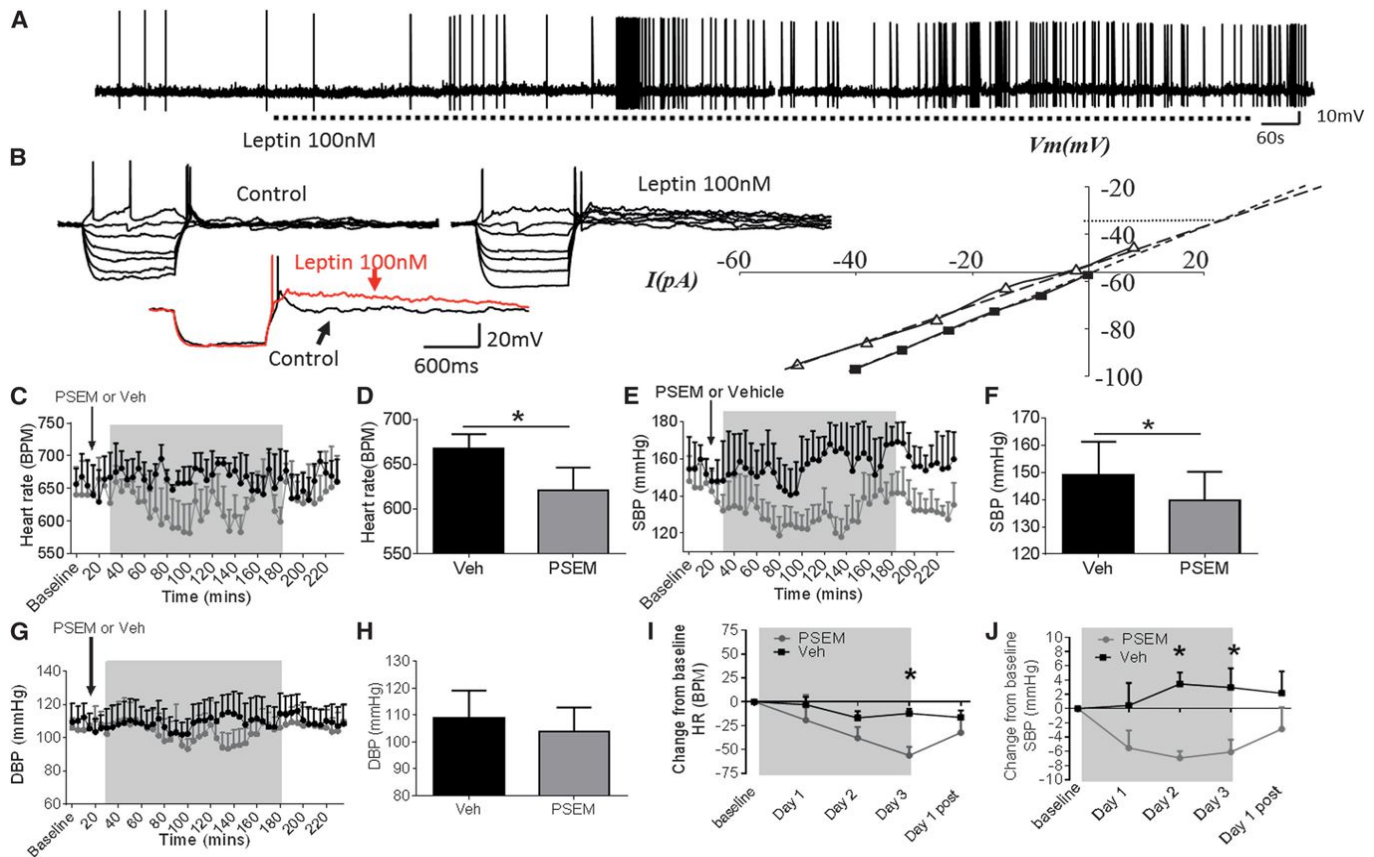


Figure 5. Leptin Excites Leptin-Receptor-Expressing Neurons in the DMH

(A) Whole-cell patch-clamp electrophysiology in LepR-Cre-YFP mice; bath application of leptin (100 nM) induced an increase in spontaneous firing rate. (B) Voltage-current relations from the cell shown in Figure S4A before and following exposure to leptin (100 nM). Note the transient rebound depolarization at the offset of the response to hyperpolarizing current injection and its prolongation in the presence of leptin, illustrated below with responses to hyperpolarizing current injection in control (black) and in the presence of leptin (red). Right, plot of data shown in (B); note the decreased slope of the voltage-current relations in the presence of leptin (100 nM) (open triangles) compared to control (closed squares), indicating a decrease in neuronal input resistance and extrapolated reversal potential around -35 mV, consistent activation of a nonselective cation conductance. (C) Averaged 5 min (data collected every 10 s) heart rate (BPM) of DIO mice injected in the DMH with inhibitor chimeric ion channel virus (pSyn::Flex-rev-PSAM L141F:GlyR-IRES-GFP) and treated i.p. with vehicle or PSEM (5 mg/kg). $n = 5$. (D) Average (20–180 min) heart rate (BPM) over treatment period of DIO mice injected in the DMH with inhibitor chimeric ion channel virus with vehicle or PSEM. Paired t test, $n = 5$. (E) Averaged 5 min (data collected every 10 s) SBP (mmHg) of DIO mice injected in the DMH with inhibitor chimeric ion channel virus and treated i.p. with vehicle or PSEM (5 mg/kg). Two-way ANOVA, Bonferroni post hoc test, $n = 5$. (F) Average (20–180 min) SBP (mmHg) in DIO mice injected in the DMH with inhibitor chimeric ion channel virus in mice treated i.p. with Veh or pharmacologically selective effector molecule (PSEM). Paired t test, $n = 5$. (G) Averaged 5 min (data collected every 10 s) DBP (mmHg) of DIO mice injected in the DMH with inhibitor chimeric ion channel virus and treated i.p. with vehicle or PSEM (5 mg/kg). Two-way ANOVA, Bonferroni post hoc test, $n = 5$. (H) Average (20–180 min) DBP (mmHg) of DIO mice treated with chimeric ion channel virus and with Veh or PSEM. Paired t test, $n = 5$. (I) Average daily HR (BPM) in DIO mice injected in the DMH with inhibitor chimeric ion channel virus treated i.p. with Veh or PSEM for 3 days. Paired t test, $n = 5$. (J) Average daily SBP (mmHg) in DIO mice treated with inhibitor chimeric ion channel virus and with Veh or PSEM for 3 days. Paired t test, $n = 5$. Values represent mean \pm SEM. * $p < 0.05$. See also Figure S4.

In Vivo Inhibition of LepR DMH-Expressing Neurons Decreases BP

In vivo, the effects of directly altering the neuronal activity of LepR-expressing DMH neurons were assessed using engineered pharmacologically selective chimeric ion channels for activating and silencing neuron activity (Magnus et al., 2011). Briefly, this technology requires the injection of a virus, which only infects and replicates in a Cre-dependent manner, into

Cre-expressing mice. The virus drives the Cre-dependent expression of a modified ion channel, containing a PSAM element. After injection (intraperitoneal injection), the otherwise inert molecule PSEM binds to an ion channel that contains a PSAM element, which opens the ion pore and allows ion flux across the cell membrane, which depolarizes or hyperpolarizes virus-infected, Cre-expressing neurons. Using male 20-week-old chow-fed lean LepR-Cre-YFP transgenic mice (Leshan et al.,

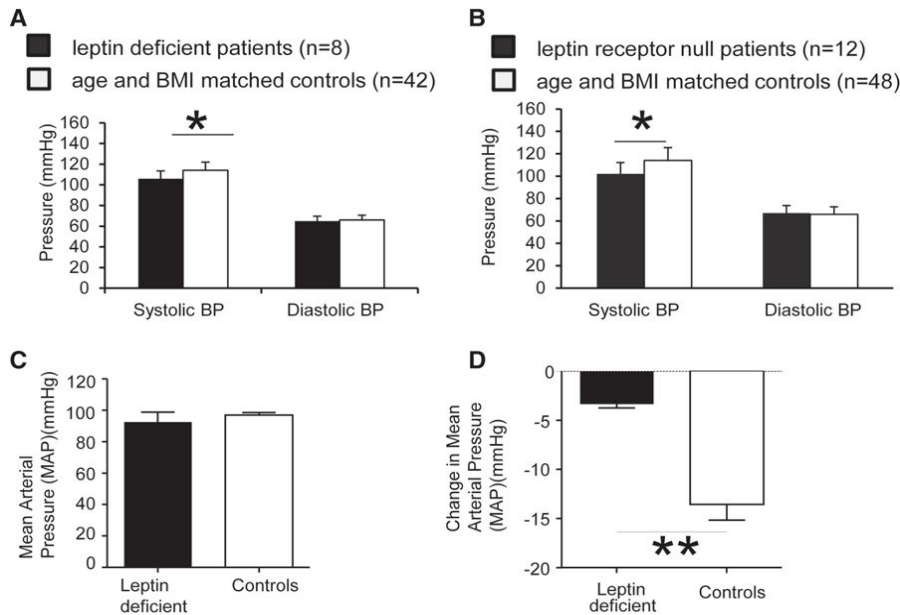


Figure 6. Blood Pressure Measurements in Humans with Leptin and Leptin Receptor Mutations

(A) Leptin-deficient children (n = 8) have significantly lower SBP (mmHg) compared to age- and weight-matched control subjects (n = 42). (B) Leptin-receptor-deficient children (n = 12) have significantly lower SBP (mmHg) compared to age- and weight-matched control subjects (n = 48). (C) Baseline pretreatment mean arterial pressure (MAP) (mmHg) of leptin-deficient adult and control subjects. n = 3. (D) Change in MAP (mmHg) after 19 weeks of weight loss treatment. n = 3. Value represent mean \pm SEM. Student's t test. *p < 0.05. See also Figure S5 and Table S1.

2012), we investigated whether activation of DMH LepR-expressing neurons could increase HR and BP. Lean LepR-Cre-YFP mice were injected bilaterally with activator virus, an engineered ionotropic serotonin receptor packaged in an AAV2 (pSyn::Flex-rev-PSAM Y115F:5HT3 HC-IRES-GFP) into the DMH. In lean mice 21 days after virus injection, twice daily intraperitoneal (i.p.) administration of the receptor ligand PSEM (i.p. 2X daily [5 mg/kg]) for 3 days significantly increased HR and BP ($p \leq 0.05$; Figures S4D–S4F) compared to vehicle-treated mice. In DIO LepR-Cre-YFP mice, an inhibitory virus, an engineered ionotropic glycine receptor packaged into an AAV2 (AAV2: pSyn::Flex-rev-PSAM L141F:GlyR-IRES-GFP) was administered bilaterally into the DMH. Compared to vehicle treated mice, PSEM acutely decreased HR (Figures 5C and 5D) and SBP (Figures 5E and 5F), but no significant change in DBP was observed (Figures 5G and 5H). Sub-chronic daily treatment with PSEM reduced HR (by 72.6 ± 5.6 bpm), SBP (by 6.1 ± 1.7 mmHg) and DBP over 3 days (Figures 5I, 5J, and S4C), effects which were reversed once PSEM treatment ceased. Treatment of mice with PSEM, prior to virus administration had no effect on HR or BP (Figures S4G and S4H). Cumulatively, these findings indicate that leptin signaling is required for the changes in BP seen in DIO and that LepR expressing neurons in the DMH are necessary and sufficient to, cause these effects. These data also support the premise that the hypertension induced by leptin in the DMH is due to leptin-induced depolarization of DMH neurons.

Human Leptin and Leptin Receptor Mutations Are Associated with Low Blood Pressure Despite Severe Obesity

Homozygous complete loss-of-function mutations in the gene encoding leptin (*LEP*) are associated with undetectable leptin levels and severe early-onset obesity in humans (Montague et al., 1997). We measured BP in the fasted rested state in eight

age- and BMI-matched controls (Figure 6A; $p < 0.05$); there was no difference in DBP. A statistically significant attenuation of SBP was also seen in severely obese children who were homozygous for complete loss-of-function mutations in the *LEPR* gene (n = 12) compared to 48 age- and BMI-matched controls (Figure 6B; $p < 0.05$) (Table S1). There were no significant differences in resting HR between the groups (data not shown). Administration of recombinant methionyl human leptin to individuals with congenital leptin deficiency leads to significant weight loss (Farooqi et al., 1999; Ozata et al., 1999). In a previously reported experimental paradigm (Galgani et al., 2010), three leptin-deficient adults were studied before and after treatment with recombinant leptin for 19 weeks, which was sufficient to cause 15.5 ± 0.5 kg weight loss. In parallel, three age- and BMI-matched controls were studied before and after a diet and exercise intervention, which achieved a comparable degree of weight loss (14.8 ± 1.76 kg). The three adult leptin-deficient individuals were found to have normal BP despite severe obesity (Figure 6C). Whereas the MAP of the obese control group decreased as expected (Figure 6D), the BP of the leptin-deficient adults did not change. No difference in HR was observed (Figure S5). Similarly, there was no statistically significant change in BP after recombinant leptin administration to leptin-deficient children (data not shown). These studies in rare individuals completely lacking leptin or functioning LEPR support the assertion that leptin is necessary for the increased BP associated with obesity in humans.

DISCUSSION

Leptin Is the Peripheral Signal Linking Weight Gain to Changes in Blood Pressure

Cumulatively, these studies demonstrate that leptin signaling is necessary for obesity-induced increased BP. We have used multiple convergent methods, including making lean mice

obese, by adding leptin systemically and by restoring LepR to the DMH of LepR-deficient mice. In all these studies, BP increased. Similarly, we have studied reduced leptin signaling in leptin- and LepR-deficient mice, neutralized circulating leptin with a systemic antibody, infused LepR antagonist ICV and intra-DMH; used shRNA to knock down LepR expression; selectively deleted LepR; and inhibited LepR-expressing neurons in the DMH of obese hypertensive mice. In all of these experiments, reducing leptin signaling reduced BP, even in the presence of obesity. Clinical studies in severely obese humans with two different forms of defective leptin signaling show that these observations are relevant to human physiology and pathophysiology.

Leptin in the DMH Regulates Cardiovascular Function in the Absence of Effects on BW

Leptin can acutely increase HR, BP, and BAT, even in anesthetized animals, presumably through activation of the SNS (Enriori et al., 2011; Marsh et al., 2003). Intracerebroventricular (ICV) leptin increases SNA to numerous organs, including the lumbar, kidney, and BAT regions (Haynes et al., 1999). Central antagonism of LepR caused a reduction in BP and HR in DIO hypertensive mice. Although early lesion studies confirm the importance of the DMH for the control of energy homeostasis, little is known about the neurochemical phenotype of the neurons present in the DMH (Elmqvist et al., 1999; Lee et al., 2012). The DMH appears to have a critical ability to control thermogenesis and is involved in elevated BAT-mediated thermogenesis found in obese mice (Enriori et al., 2011; Fan et al., 2005; Morrison et al., 2008). In the studies conducted with antagonism or knocking down of the LepR in DIO mice, we did not find significant changes in FI or BW. These results suggest that the DMH LepR-expressing neurons are not essential for leptin's effects on BW because the mice lost no additional weight when LepRs were inactivated. Although lesion studies have previously demonstrated a critical role of the DMH in the control of BW, we have shown that long-term knockdown of the LepR does not significantly affect BW, despite the significant reduction in BP and previously reported increase in BAT temperature in obesity (Elmqvist et al., 1999; Enriori et al., 2011). Additionally, acute activation or inhibition of LepR-expressing neurons did not change FI or BW.

Leptin in Humans

Contrary to our expectations, we did not observe an increase in BP with the administration of leptin to leptin-deficient patients. There are a number of possible explanations. First, the magnitude of leptin's effect on BP may be too small to be detected given the sensitivity of the tools for measurement of BP in humans and the small number of individuals studied. Notably, despite weight loss with continued leptin administration, we did not see a significant decrease in BP in these patients as would be expected in more common forms of obesity. These effects are comparable to the effects of leptin on total energy expenditure (TEE), another phenotype mediated by sympathetic tone. We have previously shown that, in contrast to weight loss in obese controls where TEE decreases, TEE does not change with leptin administration in leptin deficiency (Galgani et al., 2010).

Whether leptin increases BP (to a small degree) or leptin-deficient individuals respond differently to weight loss cannot be deduced from our data. However, our results are in line with the suggestion that the effect of leptin administration may be to prevent the reduction in BP expected in congenitally leptin-deficient individuals as they lose weight. In previous trials of leptin in common obesity, no effects on BP were observed (Heymsfield et al., 1999). Rosenbaum and Leibel have shown that controlled 10% weight loss in an experimental setting is sufficient to reduce leptin levels and is associated with decreased SNA (Rosenbaum et al., 2005).

Could Leptin Be Acting Peripherally?

It is possible that leptin could be producing some of its cardiovascular effects by acting peripherally. LepR are expressed throughout the brain, including the brainstem; additionally, LepR are expressed on cardiac myocytes, the kidney, and arteries, including the coronary arteries (Knudson et al., 2005; Purdham et al., 2004; Serradeil-Le Gal et al., 1997). We treated DIO mice peripherally with a leptin antibody and then in a separate experiment, DIO mice were treated centrally with the LepR antagonist, and in both experiments, similar responses were observed. In mice lacking the LepR, re-expression of the LepR in only the DMH caused a dramatic increase in BP and HR. The opposite effects occurred when the leptin receptors were deleted from only the DMH region in hypertensive mice. Thus, although LepR expression in peripheral regions may play a role, the results presented here clearly demonstrate a key role for the DMH LepR-expressing neurons in mediating the changes in BP in obesity.

Is Leptin Action in the DMH Mediated by the Melanocortin System?

Experiments in rodents and humans support a role for melanocortin signaling in the regulation of BP. Central administration of α -MSH increases BP and HR in wild-type mice, but not in *Mc4r*^{-/-} mice, which maintain a normal BP despite severe obesity (Kuo et al., 2003; Morgan et al., 2008; Tallam et al., 2005). In humans, heterozygous loss-of-function mutations in *MC4R* are associated with a reduced prevalence of hypertension, low SBP, lower urinary noradrenaline excretion, and reduced peripheral nerve SNA (Greenfield et al., 2009; Sayk et al., 2010). Moreover, systemic administration of a centrally active melanocortin receptor agonist acutely increased BP in obese volunteers (Greenfield et al., 2009). As such, some of leptin's effects on BP may be mediated by the melanocortin system. However, there are some indications that, in obesity, the POMC neurons in the ARH become nonresponsive to leptin (this has been termed leptin resistance), and this may limit how much the actions of leptin can be transduced by the POMC neurons in the ARH (Enriori et al., 2007; Knight et al., 2010; Münzberg et al., 2004). Previous research has implicated the POMC neurons of the ARH in the regulation of obesity-induced hypertension via activation of the IKK β /NF- κ B pathway (Purkayastha et al., 2011; Zhang et al., 2008). How POMC neurons interact with the DMH LepR-expressing neurons has not been addressed in these studies, but the DMH LepR-expressing neurons do send direct efferent projections to ARH neurons (Gautron

et al., 2010). It is possible that the LepR-expressing neurons in the DMH express Mc4r or that the DMH LepR-expressing neurons act through other neurons that express Mc4r (Liu et al., 2003). The degree to which leptin's effects on BP in obesity are dependent upon intact melanocortin signaling remains to be determined (do Carmo et al., 2011; Harlan et al., 2011).

DMH Neurons Regulate Autonomic Outflow to the Heart and Peripheral Vasculature

DMH LepR-expressing neurons are known to project to numerous brain regions, including the PVH (Elmqvist et al., 1998; Gautron et al., 2010). DMH neurons also project to other nuclei, including the Raphe pallidus nucleus (RPa) and rostral ventrolateral medulla (RVLM) (Cao and Morrison, 2003; Horiuchi et al., 2004; Simonds and Cowley, 2013). Microinjection of leptin into the DMH of anesthetized rats can induce an acute increase in HR and BP (Marsh et al., 2003). The DMH-RPa connection has already been recognized to be important in the regulation of BAT thermogenesis in response to leptin (Zhang et al., 2011). The RVLM, however, appears to have a greater control over the regulation of BP, compared to regulation of thermogenesis and HR (Horiuchi et al., 2006; Morrison et al., 2008). The neurochemical phenotype of the LepR-expressing neuron populations in the DMH is still debated (Lee et al., 2012). Further studies will be necessary to characterize the DMH circuits that contribute to leptin's effects on blood pressure and to characterize the mechanisms by which these neurons modulate sympathetic outputs to the heart and peripheral blood vessels.

Why Leptin—What about Insulin?

Although results here strongly demonstrate the role of leptin in contributing to elevated BP in obesity, a number of other hormones that change in response to weight gain could contribute. As expected, we found that, compared to lean mice, plasma insulin was elevated in all obese models. Despite this increase in plasma insulin, it was only DIO mice that exhibited significantly elevated BP, and insulin levels were not correlated with HR or BP, suggesting that insulin contributes little to the chronic elevation of BP seen in obesity. Also, insulin sensitivity as measured by hyperinsulinemic euglycemic clamps is comparable in human MC4R deficiency versus obese controls despite a lower BP and reduced urinary catecholamine excretion in MC4R-deficient subjects (Greenfield et al., 2009). Therefore, our data do not support a role for insulin in mediating obesity-induced increases in BP.

In conclusion, these observations suggest that pharmacological approaches based on the modulation of leptin's effects on specific subpopulations of neurons could represent a potentially useful therapeutic strategy for the treatment of obesity-associated hypertension and for the prevention of some obesity-associated cardiovascular disease.

EXPERIMENTAL PROCEDURES

All animal procedures were approved by Monash University animal ethics committee. All mice were housed in a controlled environment in which lights were on a 12 hr light/12 hr dark cycle; temperature and humidity remained constant. In experiments to examine the development of hypertension, 4-week-

old male C57Bl/6J mice were implanted with radiotelemetry probes (DSI, USA, model TA11PA-C10). These mice were allowed 2 weeks recovery post-surgery and, following baseline recordings, mice were split onto either chow (4.8% of fat, mouse and rat rodent chow diet, Specialty Feeds, Glen Forrest, Australia) or HFD (43% of fat, SF04-001, Specialty Feeds, Western Australia, Australia) for 20 weeks (140 days) in which recordings were taken every 13 to 15 days. After 20 weeks (140 days), the HFD fed mice were swapped onto a chow diet (4.8% of fat, SF04-001, Specialty Feeds, Western Australia, Australia) for 4 weeks. In all other experiments, male C57Bl/6J, leptin-deficient (*ob/ob*), LepR-deficient (*db/db*), LepR-deficient (LepRTB (Berglund et al., 2012), LepR flox, and LepR-Cre-YFP mice (Leshan et al., 2012) were placed on either a chow diet or HFD diet at 4 weeks of age and continuing for 20 consecutive weeks, after which mice were used for experiments. All mice were on a C57Bl/6J background. In all experiments, the animals' BW and FI were monitored daily. Specific experimental procedures for radiotelemetry, pharmacological studies, genetic manipulations, and electrophysiological studies are detailed in the Supplemental Information.

Human Studies

All human studies were conducted according to the principles outlined in the Declaration of Helsinki and after approval by local ethical committees. All individuals or their parents (for children) gave written informed consent. Systolic and diastolic BP were measured in the rested, fasted state using wrist BP monitors (OMRON Healthcare, Hamburg, Germany) in leptin-deficient ($n = 8$) and leptin-receptor-deficient children ($n = 12$) (Farooqi et al., 2002, 2007). Control subjects were recruited from the Genetics of Obesity Study (GOOS) cohort. These control subjects had been tested for mutations in leptin, leptin receptor, and MC4R. Control subjects were age and BMI matched to leptin-deficient and leptin-receptor-deficient subjects ($n = 42$ and 48 , respectively) (Wheeler et al., 2013).

Adult Studies

Leptin-deficient adult patients received pretreatment testing at the Pennington center, along with weight- and BP-matched control subjects. The leptin-deficient patients then received a 3 month treatment of recombinant leptin (Leptin [Amgen Inc., Thousand Oaks; Galgani et al., 2010]). Subjects were all provided a nutritionally balanced mixed diet during this period. Subjects then returned to the Pennington center for posttesting. Control subjects with "normal" leptin levels were administered a low-calorie diet (Galgani et al., 2010) for 9–20 weeks to cause the same weight loss and also returned to the Pennington center for posttesting. Mean \pm SEM, t test, $^{**}p < 0.01$. All subjects provided written informed consent. The study design was approved by the ethics committee of the Pennington Biomedical Research Center. Approval for leptin replacement included UCLA IRB approval (April 9, 2001) and FDA approval (IND application number 61690).

Statistics

Data are represented as mean \pm SEM, and error bars also indicate SEM. p values were calculated by either unpaired or paired two-tailed Student's t test, One-way ANOVA with Bonferroni post hoc test or two-way ANOVA with Bonferroni post hoc test. $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, five figures, and one table and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2014.10.058>.

AUTHOR CONTRIBUTIONS

The design and performance of animal experiments was conducted by S.E.S., J.T.P., J.B., R.D.B., P.J.E., D.C.S., and M.A.C. Human clinical experiments were designed and performed by E.R., F.L.G., J.L., E.H., J.M.K., S.O., and I.S.F. Experiments were assisted by contributions from R.D., A.M.A., M.G.M., K.L.G., S.M.S., and J.K.E. Data analysis was conducted by S.E.S., J.T.P., E.R., F.L.G., J.L., D.C.S., I.S.F., and M.A.C. The manuscript was prepared by S.E.S., J.T.P., D.C.S., I.S.F., and M.A.C.

ACKNOWLEDGMENTS

This work was supported by the Heart Foundation of Australia (M.A.C. and S.E.S.), the National Health and Medical Research Council of Australia (1029188 to M.A.C. and 1063955 to D.C.S.), Monash University (S.E.S.), Pfizer Australia (M.A.C.), a NORC Center grant (P30DK072476) at the Pennington Biomedical Research Center (E.R., F.L.G., and J.L.), the Leverhulme Trust (J.T.P.), the Wellcome Trust (082390/Z/07/Z) (I.S.F.), Medical Research Council (I.S.F.), NIHR Cambridge Biomedical Research Centre (I.S.F. and S.O.R.), and the Bernard Wolfe Health Neuroscience Fund (I.S.F.). We thank Professor Streamson Chua for the LepR flox mice. J.L. and I.S.F. treat patients with recombinant human leptin which is provided by BMS/AstraZeneca.

Received: May 27, 2014

Revised: August 31, 2014

Accepted: October 30, 2014

Published: December 4, 2014

REFERENCES

- Berglund, E.D., Vianna, C.R., Donato, J., Jr., Kim, M.H., Chuang, J.C., Lee, C.E., Lauzon, D.A., Lin, P., Brule, L.J., Scott, M.M., et al. (2012). Direct leptin action on POMC neurons regulates glucose homeostasis and hepatic insulin sensitivity in mice. *J. Clin. Invest.* 122, 1000–1009.
- Cao, W.H., and Morrison, S.F. (2003). Disinhibition of rostral raphe pallidus neurons increases cardiac sympathetic nerve activity and heart rate. *Brain Res.* 980, 1–10.
- Carlyle, M., Jones, O.B., Kuo, J.J., and Hall, J.E. (2002). Chronic cardiovascular and renal actions of leptin: role of adrenergic activity. *Hypertension* 39, 496–501.
- Cone, R.D. (2005). Anatomy and regulation of the central melanocortin system. *Nat. Neurosci.* 8, 571–578.
- Considine, R.V., Sinha, M.K., Heiman, M.L., Kriauciunas, A., Stephens, T.W., Nyce, M.R., Ohannesian, J.P., Marco, C.C., McKee, L.J., Bauer, T.L., et al. (1996). Serum immunoreactive-leptin concentrations in normal-weight and obese humans. *N. Engl. J. Med.* 334, 292–295.
- Cowley, M.A., Pronchuk, N., Fan, W., Dinulescu, D.M., Colmers, W.F., and Cone, R.D. (1999). Integration of NPY, AGRP, and melanocortin signals in the hypothalamic paraventricular nucleus: evidence of a cellular basis for the adipostat. *Neuron* 24, 155–163.
- Cowley, M.A., Smart, J.L., Rubinstein, M., Cerdán, M.G., Diano, S., Horvath, T.L., Cone, R.D., and Low, M.J. (2001). Leptin activates anorexigenic POMC neurons through a neural network in the arcuate nucleus. *Nature* 411, 480–484.
- do Carmo, J.M., da Silva, A.A., Cai, Z., Lin, S., Dubinina, J.H., and Hall, J.E. (2011). Control of blood pressure, appetite, and glucose by leptin in mice lacking leptin receptors in proopiomelanocortin neurons. *Hypertension* 57, 918–926.
- Dustan, H.P. (1983). Mechanisms of hypertension associated with obesity. *Ann. Intern. Med.* 98, 860–864.
- Elmqvist, J.K., Ahima, R.S., Elias, C.F., Flier, J.S., and Saper, C.B. (1998). Leptin activates distinct projections from the dorsomedial and ventromedial hypothalamic nuclei. *Proc. Natl. Acad. Sci. USA* 95, 741–746.
- Elmqvist, J.K., Elias, C.F., and Saper, C.B. (1999). From lesions to leptin: hypothalamic control of food intake and body weight. *Neuron* 22, 221–232.
- Enriori, P.J., Evans, A.E., Sinnayah, P., Jobst, E.E., Tonelli-Lemos, L., Billes, S.K., Glavas, M.M., Grayson, B.E., Perello, M., Nilini, E.A., et al. (2007). Diet-induced obesity causes severe but reversible leptin resistance in arcuate melanocortin neurons. *Cell Metab.* 5, 181–194.
- Enriori, P.J., Sinnayah, P., Simonds, S.E., Garcia Rudaz, C., and Cowley, M.A. (2011). Leptin action in the dorsomedial hypothalamus increases sympathetic tone to brown adipose tissue in spite of systemic leptin resistance. *J. Neurosci.* 31, 12189–12197.
- Esler, M., Straznicky, N., Eikelis, N., Masuo, K., Lambert, G., and Lambert, E. (2006). Mechanisms of sympathetic activation in obesity-related hypertension. *Hypertension* 48, 787–796.
- Fan, W., Voss-Andreae, A., Cao, W.H., and Morrison, S.F. (2005). Regulation of thermogenesis by the central melanocortin system. *Peptides* 26, 1800–1813.
- Farooqi, I.S., Jebb, S.A., Langmack, G., Lawrence, E., Cheetham, C.H., Prentice, A.M., Hughes, I.A., McCamish, M.A., and O'Rahilly, S. (1999). Effects of recombinant leptin therapy in a child with congenital leptin deficiency. *N. Engl. J. Med.* 341, 879–884.
- Farooqi, I.S., Matarese, G., Lord, G.M., Keogh, J.M., Lawrence, E., Agwu, C., Sanna, V., Jebb, S.A., Perna, F., Fontana, S., et al. (2002). Beneficial effects of leptin on obesity, T cell hyporesponsiveness, and neuroendocrine/metabolic dysfunction of human congenital leptin deficiency. *J. Clin. Invest.* 110, 1093–1103.
- Farooqi, I.S., Wangenstein, T., Collins, S., Kimber, W., Matarese, G., Keogh, J.M., Lank, E., Bottomley, B., Lopez-Fernandez, J., Ferraz-Amaro, I., et al. (2007). Clinical and molecular genetic spectrum of congenital deficiency of the leptin receptor. *N. Engl. J. Med.* 356, 237–247.
- Fontes, M.A., Tagawa, T., Polson, J.W., Cavanagh, S.J., and Dampney, R.A. (2001). Descending pathways mediating cardiovascular response from dorsomedial hypothalamic nucleus. *Am. J. Physiol. Heart Circ. Physiol.* 280, H2891–H2901.
- Galgani, J.E., Greenway, F.L., Caglayan, S., Wong, M.L., Licinio, J., and Ravussin, E. (2010). Leptin replacement prevents weight loss-induced metabolic adaptation in congenital leptin-deficient patients. *J. Clin. Endocrinol. Metab.* 95, 851–855.
- Gautron, L., Lazarus, M., Scott, M.M., Saper, C.B., and Elmquist, J.K. (2010). Identifying the efferent projections of leptin-responsive neurons in the dorsomedial hypothalamus using a novel conditional tracing approach. *J. Comp. Neurol.* 518, 2090–2108.
- Greenfield, J.R., Miller, J.W., Keogh, J.M., Henning, E., Satterwhite, J.H., Cameron, G.S., Astruc, B., Mayer, J.P., Brage, S., See, T.C., et al. (2009). Modulation of blood pressure by central melanocortinergic pathways. *N. Engl. J. Med.* 360, 44–52.
- Halaas, J.L., Boozer, C., Blair-West, J., Fidasein, N., Denton, D.A., and Friedman, J.M. (1997). Physiological response to long-term peripheral and central leptin infusion in lean and obese mice. *Proc. Natl. Acad. Sci. USA* 94, 8878–8883.
- Harlan, S.M., Morgan, D.A., Agassandian, K., Guo, D.F., Cassell, M.D., Sigmund, C.D., Mark, A.L., and Rahmouni, K. (2011). Ablation of the leptin receptor in the hypothalamic arcuate nucleus abrogates leptin-induced sympathetic activation. *Circ. Res.* 108, 808–812.
- Harris, R.B., Zhou, J., Redmann, S.M., Jr., Smagin, G.N., Smith, S.R., Rodgers, E., and Zachwieja, J.J. (1998). A leptin dose-response study in obese (ob/ob) and lean (+/?) mice. *Endocrinology* 139, 8–19.
- Haynes, W.G. (2000). Interaction between leptin and sympathetic nervous system in hypertension. *Curr. Hypertens. Rep.* 2, 311–318.
- Haynes, W.G., Morgan, D.A., Djalali, A., Sivitz, W.I., and Mark, A.L. (1999). Interactions between the melanocortin system and leptin in control of sympathetic nerve traffic. *Hypertension* 33, 542–547.
- Heymsfield, S.B., Greenberg, A.S., Fujioka, K., Dixon, R.M., Kushner, R., Hunt, T., Lubina, J.A., Patane, J., Self, B., Hunt, P., et al. (1999). Recombinant leptin for weight loss in obese and lean adults: a randomized, controlled, dose-escalation trial. *JAMA* 282, 1568–1575.
- Hommel, J.D., Trinko, R., Sears, R.M., Georgescu, D., Liu, Z.W., Gao, X.B., Thurmon, J.J., Marinelli, M., and DiLeone, R.J. (2006). Leptin receptor signaling in midbrain dopamine neurons regulates feeding. *Neuron* 51, 801–810.
- Horiuchi, J., McAllen, R.M., Allen, A.M., Killinger, S., Fontes, M.A., and Dampney, R.A. (2004). Descending vasomotor pathways from the dorsomedial hypothalamic nucleus: role of medullary raphe and RVLM. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 287, R824–R832.

- Horiuchi, J., McDowall, L.M., and Dampney, R.A. (2006). Differential control of cardiac and sympathetic vasomotor activity from the dorsomedial hypothalamus. *Clin. Exp. Pharmacol. Physiol.* 33, 1265–1268.
- Kassab, S., Kato, T., Wilkins, F.C., Chen, R., Hall, J.E., and Granger, J.P. (1995). Renal denervation attenuates the sodium retention and hypertension associated with obesity. *Hypertension* 25, 893–897.
- Knight, Z.A., Hannan, K.S., Greenberg, M.L., and Friedman, J.M. (2010). Hyperleptinemia is required for the development of leptin resistance. *PLoS ONE* 5, e11376.
- Knudson, J.D., Dincer, U.D., Zhang, C., Swafford, A.N., Jr., Koshida, R., Picchi, A., Focardi, M., Dick, G.M., and Tune, J.D. (2005). Leptin receptors are expressed in coronary arteries, and hyperleptinemia causes significant coronary endothelial dysfunction. *Am. J. Physiol. Heart Circ. Physiol.* 289, H48–H56.
- Konstantinides, S., Schäfer, K., Neels, J.G., Dellas, C., and Loskutoff, D.J. (2004). Inhibition of endogenous leptin protects mice from arterial and venous thrombosis. *Arterioscler. Thromb. Vasc. Biol.* 24, 2196–2201.
- Kuo, J.J., Silva, A.A., and Hall, J.E. (2003). Hypothalamic melanocortin receptors and chronic regulation of arterial pressure and renal function. *Hypertension* 41, 768–774.
- Lee, S., Bookout, A.L., Lee, C.E., Gautron, L., Harper, M.J., Elias, C.F., Lowell, B.B., and Elmquist, J.K. (2012). Laser-capture microdissection and transcriptional profiling of the dorsomedial nucleus of the hypothalamus. *J. Comp. Neurol.* 520, 3617–3632.
- Leshan, R.L., Greenwald-Yarnell, M., Patterson, C.M., Gonzalez, I.E., and Myers, M.G., Jr. (2012). Leptin action through hypothalamic nitric oxide synthase-1-expressing neurons controls energy balance. *Nat. Med.* 18, 820–823.
- Liu, H., Kishi, T., Roseberry, A.G., Cai, X., Lee, C.E., Montez, J.M., Friedman, J.M., and Elmquist, J.K. (2003). Transgenic mice expressing green fluorescent protein under the control of the melanocortin-4 receptor promoter. *J. Neurosci.* 23, 7143–7154.
- Maffei, M., Halaas, J., Ravussin, E., Pratley, R.E., Lee, G.H., Zhang, Y., Fei, H., Kim, S., Lallone, R., Ranganathan, S., et al. (1995). Leptin levels in human and rodent: measurement of plasma leptin and ob RNA in obese and weight-reduced subjects. *Nat. Med.* 1, 1155–1161.
- Magnus, C.J., Lee, P.H., Atasoy, D., Su, H.H., Looger, L.L., and Sternson, S.M. (2011). Chemical and genetic engineering of selective ion channel-ligand interactions. *Science* 333, 1292–1296.
- Mark, A.L., Shaffer, R.A., Correia, M.L., Morgan, D.A., Sigmund, C.D., and Haynes, W.G. (1999). Contrasting blood pressure effects of obesity in leptin-deficient ob/ob mice and agouti yellow obese mice. *J. Hypertens.* 17, 1949–1953.
- Marsh, A.J., Fontes, M.A., Killinger, S., Pawlak, D.B., Polson, J.W., and Dampney, R.A. (2003). Cardiovascular responses evoked by leptin acting on neurons in the ventromedial and dorsomedial hypothalamus. *Hypertension* 42, 488–493.
- Montague, C.T., Farooqi, I.S., Whitehead, J.P., Soos, M.A., Rau, H., Wareham, N.J., Sewter, C.P., Digby, J.E., Mohammed, S.N., Hurst, J.A., et al. (1997). Congenital leptin deficiency is associated with severe early-onset obesity in humans. *Nature* 387, 903–908.
- Morgan, D.A., Thedens, D.R., Weiss, R., and Rahmouni, K. (2008). Mechanisms mediating renal sympathetic activation to leptin in obesity. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 295, R1730–R1736.
- Morrison, S.F., Nakamura, K., and Madden, C.J. (2008). Central control of thermogenesis in mammals. *Exp. Physiol.* 93, 773–797.
- Münzberg, H., Flier, J.S., and Bjørbaek, C. (2004). Region-specific leptin resistance within the hypothalamus of diet-induced obese mice. *Endocrinology* 145, 4880–4889.
- Ozata, M., Ozdemir, I.C., and Licinio, J. (1999). Human leptin deficiency caused by a missense mutation: multiple endocrine defects, decreased sympathetic tone, and immune system dysfunction indicate new targets for leptin action, greater central than peripheral resistance to the effects of leptin, and spontaneous correction of leptin-mediated defects. *J. Clin. Endocrinol. Metab.* 84, 3686–3695.
- Patterson, C.M., Leshan, R.L., Jones, J.C., and Myers, M.G., Jr. (2011). Molecular mapping of mouse brain regions innervated by leptin receptor-expressing cells. *Brain Res.* 1378, 18–28.
- Poirier, P., Giles, T.D., Bray, G.A., Hong, Y., Stern, J.S., Pi-Sunyer, F.X., and Eckel, R.H.; American Heart Association; Obesity Committee of the Council on Nutrition, Physical Activity, and Metabolism (2006). Obesity and cardiovascular disease: pathophysiology, evaluation, and effect of weight loss: an update of the 1997 American Heart Association Scientific Statement on Obesity and Heart Disease from the Obesity Committee of the Council on Nutrition, Physical Activity, and Metabolism. *Circulation* 113, 898–918.
- Purdham, D.M., Zou, M.X., Rajapurohitam, V., and Karmazyn, M. (2004). Rat heart is a site of leptin production and action. *Am. J. Physiol. Heart Circ. Physiol.* 287, H2877–H2884.
- Purkayastha, S., Zhang, G., and Cai, D. (2011). Uncoupling the mechanisms of obesity and hypertension by targeting hypothalamic IKK- β and NF- κ B. *Nat. Med.* 17, 883–887.
- Rezaei-Zadeh, K., Yu, S., Jiang, Y., Laque, A., Schwartzburg, C., Morrison, C.D., Derbenev, A.V., Zsombok, A., and Münzberg, H. (2014). Leptin receptor neurons in the dorsomedial hypothalamus are key regulators of energy expenditure and body weight, but not food intake. *Mol. Metab.* 3, 681–693.
- Rosenbaum, M., Goldsmith, R., Bloomfield, D., Magnano, A., Weimer, L., Heymsfield, S., Gallagher, D., Mayer, L., Murphy, E., and Leibel, R.L. (2005). Low-dose leptin reverses skeletal muscle, autonomic, and neuroendocrine adaptations to maintenance of reduced weight. *J. Clin. Invest.* 115, 3579–3586.
- Sayk, F., Heutling, D., Dodt, C., Iwen, K.A., Wellhoner, J.P., Scherag, S., Hinney, A., Hebebrand, J., and Lehnert, H. (2010). Sympathetic function in human carriers of melanocortin-4 receptor gene mutations. *J. Clin. Endocrinol. Metab.* 95, 1998–2002.
- Scott, M.M., Lachey, J.L., Sternson, S.M., Lee, C.E., Elias, C.F., Friedman, J.M., and Elmquist, J.K. (2009). Leptin targets in the mouse brain. *J. Comp. Neurol.* 514, 518–532.
- Serradeil-Le Gal, C., Raufaste, D., Brossard, G., Pouzet, B., Marty, E., Maffrand, J.P., and Le Fur, G. (1997). Characterization and localization of leptin receptors in the rat kidney. *FEBS Lett.* 404, 185–191.
- Simonds, S.E., and Cowley, M.A. (2013). Hypertension in obesity: is leptin the culprit? *Trends Neurosci.* 36, 121–132.
- Tallam, L.S., Stec, D.E., Willis, M.A., da Silva, A.A., and Hall, J.E. (2005). Melanocortin-4 receptor-deficient mice are not hypertensive or salt-sensitive despite obesity, hyperinsulinemia, and hyperleptinemia. *Hypertension* 46, 326–332.
- Wheeler, E., Huang, N., Bochukova, E.G., Keogh, J.M., Lindsay, S., Garg, S., Henning, E., Blackburn, H., Loos, R.J., Wareham, N.J., et al. (2013). Genome-wide SNP and CNV analysis identifies common and low-frequency variants associated with severe early-onset obesity. *Nat. Genet.* 45, 513–517.
- Zhang, Y., Proenca, R., Maffei, M., Barone, M., Leopold, L., and Friedman, J.M. (1994). Positional cloning of the mouse obese gene and its human homologue. *Nature* 372, 425–432.
- Zhang, X., Zhang, G., Zhang, H., Karin, M., Bai, H., and Cai, D. (2008). Hypothalamic IKK β /NF- κ B and ER stress link overnutrition to energy imbalance and obesity. *Cell* 135, 61–73.
- Zhang, Y., Kerman, I.A., Laque, A., Nguyen, P., Faouzi, M., Louis, G.W., Jones, J.C., Rhodes, C., and Münzberg, H. (2011). Leptin-receptor-expressing neurons in the dorsomedial hypothalamus and median preoptic area regulate sympathetic brown adipose tissue circuits. *J. Neurosci.* 31, 1873–1884.

Identification of Spinal Circuits Transmitting and Gating Mechanical Pain

Bo Duan,^{1,6} Longzhen Cheng,^{1,5,6} Steeve Bourane,² Olivier Britz,² Christopher Padilla,² Lidia Garcia-Campmany,² Michael Krashes,^{3,7} Wendy Knowlton,¹ Tomoko Velasquez,² Xiangyu Ren,¹ Sarah E. Ross,⁴ Bradford B. Lowell,³ Yun Wang,⁵ Martyn Goulding,^{2,*} and Qiufu Ma^{1,*}

¹Dana-Farber Cancer Institute and Department of Neurobiology, Harvard Medical School, 1 Jimmy Fund Way, Boston, MA 02115, USA

²Molecular Neurobiology Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA

³Division of Endocrinology, Department of Medicine, Beth Israel Deaconess Medical Center and Harvard Medical School, 99 Brookline Avenue, Boston, MA 02215, USA

⁴Department of Neurobiology and University of Pittsburgh Pain Center, University of Pittsburgh, 200 Lothrop Street, Pittsburgh, PA 15213, USA

⁵Institute of Brain Science and State Key Laboratory of Medical Neurobiology, Fudan University, Shanghai 200032, China

⁶Co-first authors

⁷Present address: National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD 20892, USA and National Institute on Drug Abuse, National Institutes of Health, Baltimore, MD 21224, USA

*Correspondence: goulding@salk.edu (M.G.), Qiufu_Ma@dfci.harvard.edu (Q.M.)

<http://dx.doi.org/10.1016/j.cell.2014.11.003>

SUMMARY

Pain information processing in the spinal cord has been postulated to rely on nociceptive transmission (T) neurons receiving inputs from nociceptors and A β mechanoreceptors, with A β inputs gated through feed-forward activation of spinal inhibitory neurons (INs). Here, we used intersectional genetic manipulations to identify these critical components of pain transduction. Marking and ablating six populations of spinal excitatory and inhibitory neurons, coupled with behavioral and electrophysiological analysis, showed that excitatory neurons expressing somatostatin (SOM) include T-type cells, whose ablation causes loss of mechanical pain. Inhibitory neurons marked by the expression of dynorphin (Dyn) represent INs, which are necessary to gate A β fibers from activating SOM⁺ neurons to evoke pain. Therefore, peripheral mechanical nociceptors and A β mechanoreceptors, together with spinal SOM⁺ excitatory and Dyn⁺ inhibitory neurons, form a microcircuit that transmits and gates mechanical pain.

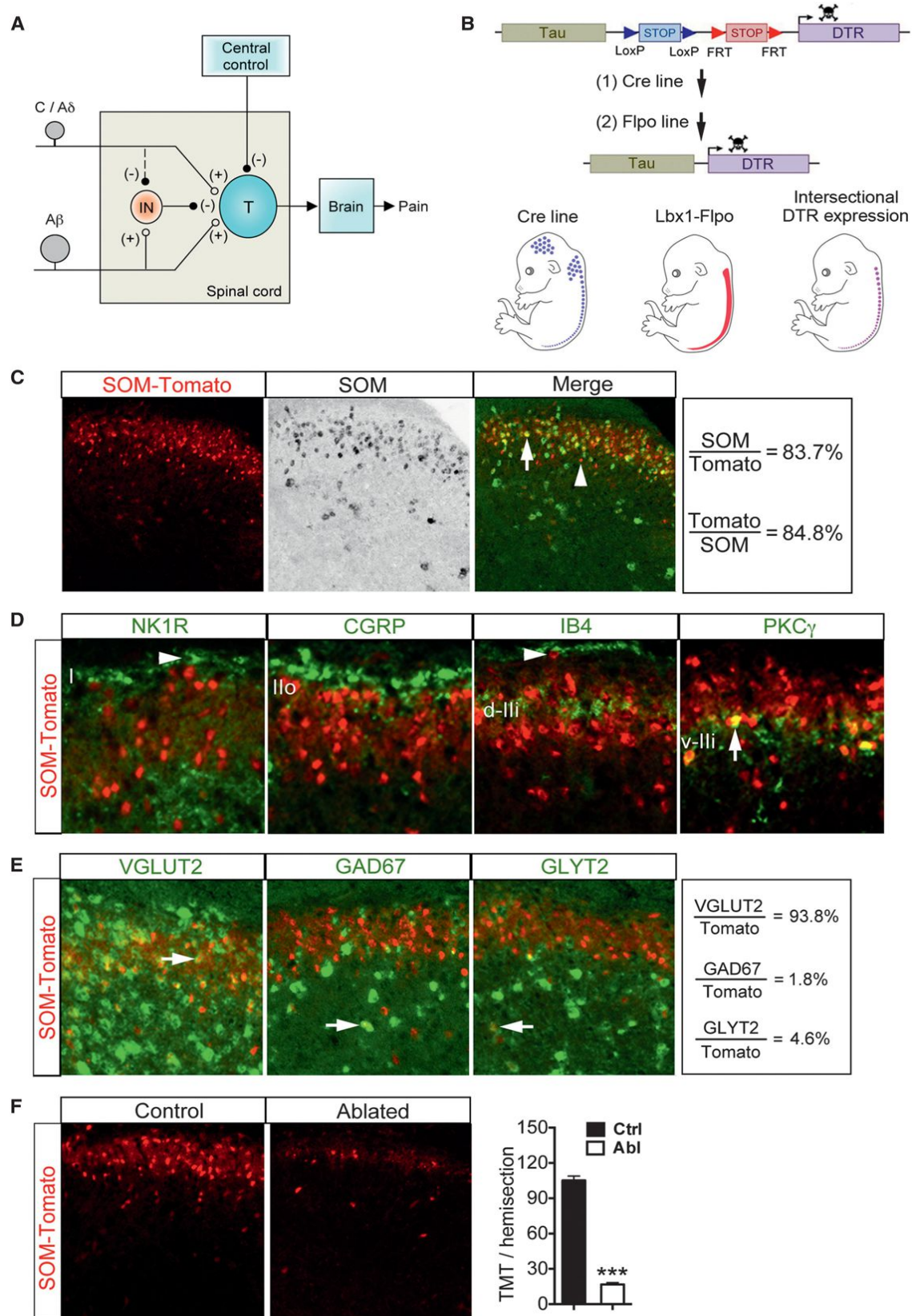
INTRODUCTION

The dorsal spinal cord is the integrative center that processes and transmits a variety of somatic sensory modalities, such as pain, itch, cold, warmth, and touch. In the past century, two dominant theories, specificity versus pattern, have been proposed to explain how pain modalities are encoded. In the late 1960s, Perl and colleagues identified nociceptors in the dorsal root ganglia (DRG) and nociceptive relay neurons in the dorsal spinal cord, lending support for the existence of pain-specific circuits (Bessou and Perl, 1969; Burgess and Perl, 1967; Chris-

tensen and Perl, 1970). Meanwhile, the pattern theory argues that processing of pain-related information can be modulated by brain states and by inputs from other types of sensory fibers (Head, 1905; Melzack and Wall, 1982; Noordenbos, 1987). In particular, the gate control theory of pain, proposed by Melzack and Wall in 1965 and revised in 1978, argues that spinal nociceptive transmission (T) neurons also receive inputs from low-threshold A β mechanoreceptors, but this input is gated by feed-forward activation of inhibitory neurons (INs) located in the substantia gelatinosa (lamina II) of the dorsal horn (Melzack and Wall, 1965; Wall, 1978) (Figure 1A).

Nearly 50 years later, numerous studies tried to test the key argument of the gate control theory of pain (Braz et al., 2014; Mendell, 2014). First, this theory correctly predicts that disinhibition could be a reason for the manifestation of mechanical allodynia or pain evoked by innocuous mechanical stimuli (Prescott et al., 2014; Price et al., 2009; Sandkühler, 2009; Zeilhofer et al., 2012). Second, electrophysiological studies have revealed the existence of a polysynaptic excitatory circuit that links A β fibers from lamina III to lamina I ascending projection neurons (Baba et al., 2003; Lu et al., 2013; Miraucourt et al., 2007; Torsney and MacDermott, 2006).

Despite this progress, precise identities of spinal neurons that transmit and gate pain-related information remain unknown (Braz et al., 2014; Prescott et al., 2014). Dorsal horn excitatory and inhibitory neurons are extremely heterogeneous, as indicated by distinct molecular markers, firing patterns, and morphologies (Ribeiro-da-Silva and De Koninck, 2008; Todd, 2010). One effective way to identify the spinal neurons required to process somatic sensory information has been the use of saporin-conjugated peptides to ablate spinal neurons expressing specific peptide receptors (Carstens et al., 2010; Mantyh et al., 1997; Mishra and Hoon, 2013; Sun et al., 2009). However, this approach has a potential complication, which is that intrathecal injection of a saporin-conjugated peptide might ablate central terminals originating from primary sensory neurons that also express the receptor for this particular peptide.



(legend on next page)

Thus, to date, it is still not known what spinal excitatory neurons are required to sense specific pain submodalities, such as thermal versus mechanical pain. Nor is it known about the identities of the inhibitory neurons that gate pain-related information.

Here, we have designed an intersectional genetic strategy (Dymecki and Kim, 2007) that allows us to specifically mark and ablate a cohort of molecularly defined subpopulations of spinal excitatory or inhibitory neurons. Subsequent behavioral and electrophysiological studies have identified two populations of spinal neurons, the somatostatin (SOM) lineage excitatory neurons and the dynorphin (Dyn) lineage inhibitory neurons, as components of the spinal circuits that transmit and gate mechanical pain.

RESULTS

Intersectional Genetic Ablation of Dorsal Spinal Excitatory and Inhibitory Neurons

To map spinal circuits processing somatic sensory information, we used an intersectional genetic strategy to ablate individual populations of spinal excitatory and inhibitory neurons. To do this, three sets of mouse lines are involved (Figure 1B). The first one is the intersectional *Tau^{loxP-STOP-loxP-FRT-STOP-FRT-DTR}* (or *Tau^{DTR/+}*) mice, in which the human diphtheria toxin receptor (DTR) gene (Saito et al., 2001) is driven from the pan neuronal *Tau* promoter (Figure 1B). The DTR expression is, however, not activated until after removal of two STOP cassettes by the Cre and flippase (Flpo) DNA recombinases. The second line is *Lbx1^{Flpo/+}*, in which *Flpo* is driven from the *Lbx1* promoter. Importantly, *Lbx1-Flpo* drove reporter expression only in neurons derived from the dorsal spinal cord and the dorsal hindbrain (Figure S1A and S1B available online). Moreover, the *Lbx1* lineage neurons include all the inhibitory neurons located in the dorsal horn (Gross et al., 2002; Müller et al., 2002) and excitatory neurons required to sense pain and itch (Xu et al., 2013). The third set of mouse lines includes various *Cre* lines. By crossing these three sets of mouse lines together (*Tau^{DTR/+}*, *Lbx1^{Flpo/+}* and *Cre* mice), only spinal neurons that express both Cre and Flpo will remove both STOP cassettes and activate DTR expression (Figure 1B). Upon diphtheria toxin (DTX) injection, these DTR-expressing spinal neurons can be ablated selectively.

In total, we ablated and analyzed six lineages of spinal neurons, with a specific goal of identifying neurons involved with transmission and/or gate control of mechanical pain. Three lineages represent predominantly excitatory neurons marked by *Cre* driven from the somatostatin gene (*SOM-Cre*), the calbindin 2/calretinin gene (*Calb2-Cre*), or the preprotachykinin 2 gene (*Tac2-Cre*) (Mar et al., 2012; Taniguchi et al., 2011). We found

that only SOM lineage excitatory neurons are required to sense mechanical pain (see below). Three other lineages of spinal neurons are mainly inhibitory and are marked by *Cre* driven from the preprodynorphin gene (*Pdyn-IRES-Cre*, referred here to as *Dyn-Cre*) (Krashes et al., 2014), the neuropeptide Y gene (*NPY-Cre*), or the choline acetyltransferase gene (*ChAT-Cre*) (Rossi et al., 2011). We found that only the Dyn lineage inhibitory neurons are required to gate mechanical pain. In the remaining text, we will present evidence that SOM excitatory neurons and Dyn inhibitory neurons form a circuit for the gate control of mechanical pain.

Genetic Marking of Spinal SOM Lineage Excitatory Neurons

To mark SOM lineage neurons with Tomato expression, we crossed *SOM^{Cre/+}* mice (Taniguchi et al., 2011) with *ROSA26^{CAG-loxP-STOP-loxP-tdTomato}* reporter mice, simplified as *ROSA26^{tdTomato/+}* mice (Madisen et al., 2010), with resulting double heterozygous mice referred to as *SOM-Tomato*. Double staining shows that 84% (1,022/1,221) of Tomato⁺ neurons exhibit detectable SOM mRNA, and 85% (1,022/1,205) of SOM mRNA⁺ neurons coexpress Tomato (Figure 1C), indicating that *SOM-Cre* faithfully marks most SOM⁺ neurons. The 16% of *SOM-Tomato*⁺ neurons without detectable SOM mRNA could represent neurons with transient SOM expression.

We next determined laminar distribution of *SOM-Tomato*⁺ neurons. NK1R expression marks a large fraction of ascending projection neurons located in dorsal horn lamina I (Todd, 2010). *SOM-Tomato*⁺ neurons are located mainly ventral to NK1R⁺ neurons, with almost none (0/98) of lamina I neurons with high NK1R expression coexpressing *SOM-Tomato* (Figure 1D). Lamina II is subdivided into three sublayers. The outer layer (II_o) is innervated by CGRP⁺ peptidergic DRG neurons. The dorsal inner layer (d-II_i) is innervated by DRG neurons labeled by isolectin B4 (IB4), and the ventral inner layer (v-II_i) is partly defined by interneurons that express protein kinase C γ (PKC γ) (Braz et al., 2014; Todd, 2010). *SOM-Tomato*⁺ neurons are intermingled with CGRP⁺ terminals in II_o and with IB4⁺ terminals in d-II_i. The ventral limit of dense *SOM-Tomato*⁺ neurons matches with dense PKC γ ⁺ neurons in v-II_i, and a subset of PKC γ ⁺ neurons coexpress Tomato (Figure 1D). Double staining with NeuN, which marks most, but not all, dorsal horn neurons, shows that *SOM-Tomato*⁺ neurons represent 7% (25/348) and 37% (711/1926) of NeuN⁺ neurons in lamina I and lamina II, respectively (Figure S1C). Thus, *SOM-Tomato*⁺ neurons are confined mainly to lamina II, but also scattered in laminae I and III-V (Figure 1C and 1D).

Figure 1. Intersectional Ablation of SOM lineage Neurons in Spinal Dorsal Horn

(A) Schematic showing the modified gate control theory of pain. “T” represents a spinal pain transmission neuron. “IN”: an inhibitory neuron. “(+)” and “(–)” represent excitatory and inhibitory inputs, respectively. The dashed line from C/A δ to IN indicates that C/A δ fibers might activate an unknown pathway to silence IN activity, but this pathway and descending modulation from brain were not studied here.

(B) Schematic showing strategy of intersectional ablation in the dorsal spinal cord. “DTR”: diphtheria toxin receptor.

(C and D) Double staining of Tomato with SOM mRNA (C) or with other markers (D), on sagittal (NK1R) or transverse (others) lumbar spinal sections of adult *SOM-Tomato* mice. Arrows indicate colocalization. Arrowheads indicate lamina I neurons with singular expression of NK1R or Tomato.

(E) Double labeling of Tomato and with indicated mRNAs. Arrows indicate colocalization.

(F) Ablation of *SOM-Tomato*⁺ neurons in lumbar dorsal spinal cord (105 \pm 4 in control [“Ctrl”] group versus 17 \pm 2 in ablated [“Abl”] group, n = 15–17 hemisections from three mice per group; p < 0.001, Student’s unpaired t test). Data are represented as mean \pm SEM. See also Figure S1.

Regarding neurotransmitter phenotypes, 94% (1,139/1,214) of Tomato⁺ neurons express the vesicular glutamate transporter VGLUT2 (Figure 1E), a marker for glutamatergic excitatory neurons (Freneau et al., 2004). SOM-Tomato⁺ neurons represent 19% (62/329) and 53% (884/1,677) of VGLUT2⁺ neurons in lamina I and lamina II, respectively. Only ~2% (21/1,277) of SOM-Tomato⁺ neurons express the GABAergic inhibitory neuron marker GAD67 (Zeilhofer et al., 2012). Additionally, about 5% (65/1,417) of SOM-Tomato⁺ neurons express the glycinergic inhibitory neuron marker GLYT2 (Zeilhofer et al., 2012), which are scattered mainly in laminae III–V (Figure 1E). Thus, a majority of SOM-Tomato⁺ neurons are excitatory, consistent with previous reports (Yasaka et al., 2010). These SOM-Tomato⁺ excitatory neurons are heterogeneous, with distinct firing patterns and morphologies (Figure S1D and S1E).

Ablation of SOM Lineage Neurons Leads to Loss of Acute Mechanical Pain

To ablate SOM lineage neurons, we crossed SOM^{Cre/+} mice with *Tau^{DTR/+}* and *Lbx1^{Flpo/+}* mice. To monitor ablation efficacy, they were further crossed with *ROSA^{Tomato/+}* reporter mice to mark SOM lineage neurons. The resulting *Tau^{DTR/+};ROSA^{Tomato/+};Lbx1^{Flpo/+};SOM^{Cre/+}* quadruple heterozygous mice were injected twice with DTX, and these mice are referred to as SOM ablated (Abl) mice. Four weeks after DTX injection, SOM-Tomato⁺ neurons were ablated in lumbar dorsal spinal cord by 82% (Figure 1F) and in the hindbrain spinal trigeminal nucleus (Sp5) (Figure S2A), but not in DRGs or other brain regions (Figure S2A).

We next performed behavioral analyses in SOM Abl mice, using littermates that lacked DTR expression but received the same DTX injections as controls. We found that ablation of SOM neurons did not affect sensorimotor coordination or the senses of innocuous touch, heat or cold (Figures S2B–S2H). In contrast, mechanical pain was markedly impaired. We first used von Frey filaments to deliver punctate mechanical stimuli onto the plantar hindpaw. SOM Abl mice showed no response at all, even with the maximal strength (2.56 g for the cutoff) used by the up-down method (Figure 2A) (Chaplan et al., 1994), and this loss is further confirmed by measuring withdrawal percentages to repeated von Frey fiber stimulation (Figure 2B). We next performed the pinprick test onto the hindpaw plantar surface, which evoked withdrawal responses in control mice, but not in SOM Abl mice (Figure 2C). We finally performed the pinch test, by placing an alligator clamp onto the hindpaw plantar surface, and measured licking responses. Licking behavior involves supraspinal processing of noxious sensory information and is considered to be a readout of feeling pain (Wang et al., 2013). The duration of licking is greatly reduced in SOM Abl mice (Figure 2D), further suggesting impairment of mechanical pain. In contrast, neither Tac2 nor Calb2 lineage neurons play major roles in sensing mechanical pain, except for a loss of sensing light punctate mechanical stimuli in Calb2 Abl mice (Figures S3, S4, and S5).

C and A δ Fiber Inputs onto SOM-Tomato⁺ Neurons in Lamina II

We next examined sensory afferent inputs onto SOM-Tomato⁺ neurons. We first performed dorsal root compound action potential recordings to determine the electric stimulation intensities

required to activate A β , A δ , and C fibers. In total, six mice at P23–P26 were used. The thresholds for A β , A δ , and C fibers, as indicated by fast, medium, and slow conduction velocities, are 12–16 μ A, 30–35 μ A, and 150–300 μ A, respectively (Figures S1F and S1G). Accordingly, the intensity ranges used in this study for different fibers are: \leq 25 μ A for A β , 30–100 μ A for A δ , and 150–500 μ A for C fibers.

We next prepared spinal cord slices with attached dorsal root, and whole-cell patch configuration was used to record synaptic inputs onto SOM-Tomato⁺ neurons directly visualized under a fluorescent microscope. Three recording conditions were used. First, to detect both large and small evoked excitatory postsynaptic currents (eEPSCs), we held the membrane potential at -70 mV to minimize evoked inhibitory postsynaptic currents (eIPSCs) (Yoshimura and Nishi, 1995). High-frequency stimulation was then used to determine monosynaptic inputs, as indicated by one-on-one responses (Baba et al., 2003; Lu and Perl, 2005; Torsney and MacDermott, 2006). It should be noted that a lack of one-on-one responses to high-frequency stimulation is often used to indicate polysynaptic inputs (Baba et al., 2003; Torsney and MacDermott, 2006) but could also indicate monosynaptic inputs with feed-forward inhibition (Bruno, 2011). Second, by holding the membrane potential at -45 mV, both eEPSCs and eIPSCs can be recorded. Third, we used current clamp mode to record evoked excitatory postsynaptic potentials (eEPSPs) to determine whether the stimulation drove action potential (AP) firing at the resting membrane potential.

We first recorded A δ and C fiber inputs that are known to include nociceptors (A β inputs will be described in the next section). In total, 41 SOM-Tomato⁺ neurons from eight mice at P23–P30 were recorded. We found that 100% of SOM-Tomato⁺ neurons in lamina II receive C fiber inputs, 89% (17/19) of them receive monosynaptic inputs indicated by one-on-one responses to high-frequency stimulations at 1 Hz, and 50% (8/16) of them generated AP output (Figure 2E–2G). A majority of SOM-Tomato⁺ neurons located at the lamina II–III border (18/22) also received C fiber inputs, 33% of which generated AP output (Figure 2E–2G). Finally, over half of SOM-Tomato⁺ neurons received A δ fiber inputs, but only 17%–21% of them generated AP outputs (Figure 2G).

Earlier in vivo extracellular recordings showed that spinal neurons located in II_o and d-II_i predominantly receive nociceptive inputs (Bennett et al., 1980; Cervero et al., 1979; Cervero et al., 1976; Christensen and Perl, 1970; Kumazawa and Perl, 1978; Light et al., 1979). Given the loss of acute mechanical pain in SOM Abl mice, C and/or A δ neurons that form synaptic connections with SOM-Tomato⁺ neurons in II_o and d-II_i likely represent mechanical nociceptors. These lamina II SOM-Tomato⁺ neurons may be directly or indirectly connected to projection neurons enriched in lamina I and lamina V (Todd, 2010) (summarized in Figure 2H). SOM-Tomato⁺ neurons in v-II_i and at II–III border might receive inputs from low threshold C/A δ mechanoreceptors (Abraira and Ginty, 2013).

A β Inputs onto SOM-Tomato⁺ Neurons in the Spinal Dorsal Horn

According to the gate control theory, spinal pain transmission neurons also receive inputs from low-threshold A β

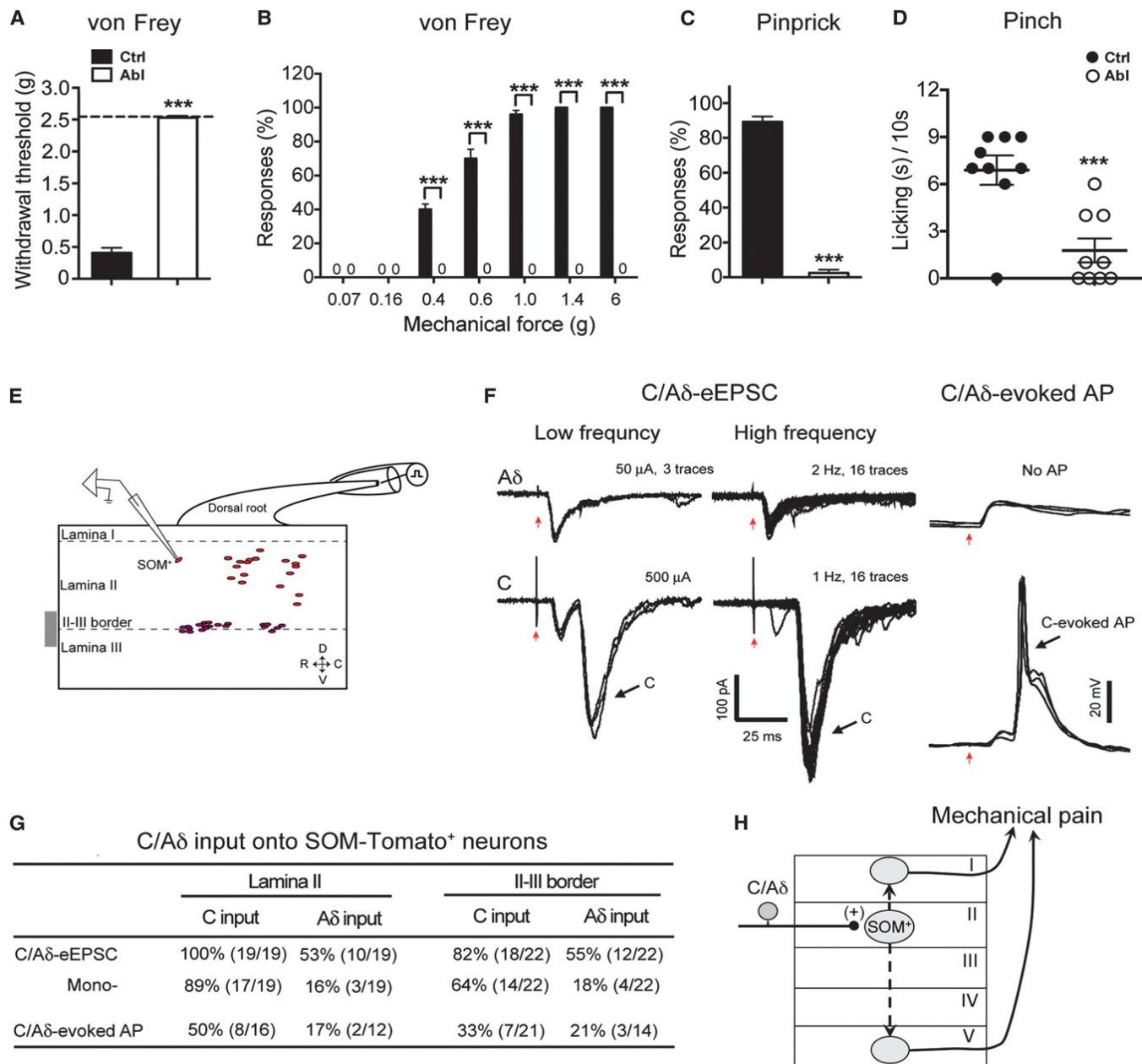


Figure 2. Loss of Acute Mechanical Pain in SOM Abl Mice and C/Aδ Inputs onto SOM-Tomato⁺ Neurons

(A) Increase of withdrawal thresholds to von Frey fiber stimulation in SOM Abl mice ("Abl") by up-down method ($n = 13$ in control ["Ctrl"] group, $n = 11$ in Abl group; *** $p < 0.001$, Student's unpaired t test).

(B) Reduced withdrawal percentages in SOM Abl mice in response to von Frey filaments ($n = 5$ in each group; *** $p < 0.001$, Student's unpaired t test).

(C) Lost response to pinprick stimulation in SOM Abl mice ($n = 13$ in Ctrl group, $n = 11$ in Abl group; *** $p < 0.001$, Student's unpaired t test).

(D) Greatly attenuated licking/flinching response to pinching in Abl mice ($n = 9$ in Ctrl group, $n = 9$ in Abl group; *** $p < 0.001$, Student's unpaired t test).

(E) Schematic showing relative positions of recorded SOM-Tomato⁺ neurons.

(F) Typical traces of C/Aδ-evoked EPSCs and APs showing C/Aδ-fiber inputs onto SOM-Tomato⁺ neurons. Red arrows indicate stimulation artifacts.

(G) The table is a summary of inputs in 41 recorded SOM-Tomato⁺ neurons from eight mice.

(H) Schematic showing that SOM-Tomato⁺ neurons in lamina II receive mono-C/Aδ input and transmit noxious signaling to lamina I and/or V pain output neurons, either directly or indirectly (dashed arrows). Data are represented as mean \pm SEM. See also Figure S1, S2, S3, S4, and S5.

mechanoreceptors that normally terminate in laminae III-V (Figure 1A). To assess A β inputs onto SOM-Tomato⁺ neurons, the dorsal root was stimulated at the A β intensity range ($\leq 25 \mu A$)

and in total, 47 SOM-Tomato⁺ neurons from nine mice at P23-P30 were recorded. We identified three types of SOM-Tomato⁺ neurons. Both type 1 and type 2 neurons are located at the II-III

border, and type 1 cells (4/18) receive monosynaptic A β inputs with AP output, and type 2 cells (14/18) receive fast A β inputs with feed-forward inhibition and do not generate A β -evoked APs under normal ACSF recording conditions. In the presence of bicuculline and strychnine, type 2 neurons can, however, generate A β -evoked fast APs, and a subset of them fire slow APs as well (Figure 3A and 3B), indicating that A β inputs onto type 2 neurons are gated by bicuculline-sensitive GABA_A and/or strychnine-sensitive glycine receptors. Five of 14 type 2 neurons show relatively large A β -evoked EPSCs and generated one-on-one responses to high-frequency stimulation, indicating monosynaptic inputs (Figure 3A).

Type 3 neurons represent most SOM-Tomato⁺ neurons within lamina II. Like type 2 neurons, they receive fast (latency < 10 ms) A β inputs without AP output (Figure 3A). Importantly, the amplitudes of A β -evoked fast EPSCs do not increase in the presence of bicuculline and strychnine (Figure 3B), thereby distinguishing them from type 2 neurons. Bicuculline and strychnine treatment did, however, result in long-lasting A β -evoked EPSCs with a slow onset (with latency \geq 10 ms) and multiple APs (Figure 3B). As described below, SOM-Tomato⁺ neurons are required to transmit A β inputs onto lamina I and II neurons. Thus, type 3 neurons receive a fast A β input (either directly or indirectly via type 1 neurons) that is gated through a mechanism insensitive to bicuculline and strychnine, and a slow A β input (possibly via type 2 neurons) that is gated by bicuculline/strychnine-sensitive feed-forward inhibition (summarized in Figure 3C).

Loss of A β Inputs onto Superficial Dorsal Horn Neurons in SOM Abl Mice

We next recorded A β inputs in spinal cord slices prepared from control and SOM Abl mice with and without the presence of bicuculline/strychnine. In total, 16 control mice (P23–P30) and 9 ablated mice (P26–P30; 7–12 days after the first DTX injection) were used. In II–III border neurons from control mice, A β fiber stimulation under the normal ACSF recording conditions drove AP firing in 5% (2/38) of recorded neurons (see below, Figure 7A), and this percentage increased to 74% (14/19) in the presence of bicuculline and strychnine, with A β stimulation evoking both fast and slow AP firing (Figure 4A). In other words, A β inputs onto 69% (74%–5%) of II–III border neurons are gated through bicuculline-sensitive GABA_A and/or strychnine-sensitive glycine receptors. In SOM Abl mice, none (0/24) of the II–III border neurons could generate A β -evoked APs (Figure 4A). In I/II_o neurons from control mice, A β fiber stimulation under normal ACSF recording conditions drove AP firing in 7% (5/69) of neurons (see below, Figure 7A), and this percentage was increased to 85% (11/13) in the presence of bicuculline and strychnine (Figure 4A), with A β stimulation evoking slow, but no fast, AP firing. In SOM Abl mice, only 4% (1/23) of I/II_o neurons could generate AP firing under the same disinhibition conditions. Collectively, these results show that SOM-Tomato⁺ neurons are required to relay A β inputs from the lamina II–III border to lamina I (Figure 4B).

Loss of Mechanical Allodynia in SOM Abl Mice

A hallmark of inflammatory and neuropathic pain is the manifestation of allodynia or pain evoked by low-threshold mechanical stimuli (Zeilhofer et al., 2012). With the loss of A β inputs onto su-

perficial dorsal horn neurons of SOM Abl mice, we next asked whether mechanical allodynia was compromised. To model inflammatory pain, complete Freund's adjuvant (CFA) was injected into the plantar pad of the hindpaw and, to assess neuropathic pain, we used the spared nerve injury (SNI) model (Decosterd and Woolf, 2000). Two types of mechanical allodynia, static and dynamic, are observed in human patients (Campbell and Meyer, 2006). Static allodynia is evoked by punctate stimuli and measured by the von Frey assay. Dynamic allodynia is evoked by movement across the skin and is mediated by A β fibers (Campbell et al., 1988; Koltzenburg et al., 1992). In mice, dynamic allodynia was measured by stroking the hindpaw plantar surface with a soft paintbrush, using the scoring system developed by Dr. Enrique José Cobos (personal communication). The typical response of naive mice to dynamic stimuli is briefly lifting the paw and walking away. This response was used for the touch assay described in Figure S2C, but for allodynia measurement, this baseline response was scored as 0. After inflammation and nerve lesions, dynamic allodynia is scored as follows: 1 for sustained lifting of the paw toward the body, 2 for strong lateral lifting above the level of the body, and 3 for flinching or licking of the affected paw. Strikingly, both static and dynamic allodynia were abolished or greatly reduced in SOM Abl mice (Figure 4C and Figure S2I), without affecting heat hyperalgesia (Figure S2J). In contrast, nerve lesion-induced mechanical allodynia is unaffected in either Tac2 Abl mice (Figure S3I) or Calb2 Abl mice (Figure S5K).

Genetic Marking and Ablation of Dyn-Expressing Spinal Inhibitory Neurons

The above studies show that A β input onto most SOM-Tomato⁺ neurons is gated by feed-forward inhibition. In the remaining sections, we will present evidence supporting the model that the Dyn lineage inhibitory neurons marked by *Dyn-Cre* act to gate mechanical pain.

We crossed *Dyn^{Cre/+}* mice with the *Rosa^{Tomato/+}* reporter to mark Dyn lineage neurons. Seventy-four percent (147/199) of Dyn-Tomato⁺ neurons exhibited detectable Dyn mRNA, and 95% (173/183) of neurons with detectable Dyn mRNA coexpressed Tomato (Figure 5A). Thus, Dyn-Cre marks most neurons with persistent Dyn expression, and a small number of neurons that likely express Dyn transiently. Dyn-Tomato⁺ neurons are located mainly in laminae I and II, and minorly in laminae III–V (Figure 5A). Eighty-six percent (151/175) of them are GAD67⁺ GABAergic inhibitory neurons, but only a small subset of GAD67⁺ neurons coexpress Dyn-Tomato (Figure 5B). Twenty-eight percent (51/189) of Dyn-Tomato⁺ neurons are GLYT2⁺ glycinergic inhibitory neurons, and they are located close to the II–III border (Figure 5B). Only 12% (24/202) are VGLUT2⁺ glutamatergic neurons (Figure 5B). The predominant association with inhibitory neurons is consistent with previous reports (Sardell et al., 2011). Consistently, half of the Dyn-Tomato⁺ neurons exhibit tonic firing (Figure S7A), a pattern shared by many inhibitory interneurons (Yasaka et al., 2010).

To examine the function of Dyn lineage neurons, we generated Dyn Abl mice using the same method we used to generate SOM Abl mice. The vast majority of Dyn-Tomato⁺ inhibitory neurons marked by GAD67 or GLYT2 were ablated in the dorsal horn



(C) Schematic showing A β inputs into types 1–3 of SOM-Tomato⁺ neurons. “IN”: inhibitory neurons. See also Figure S1.

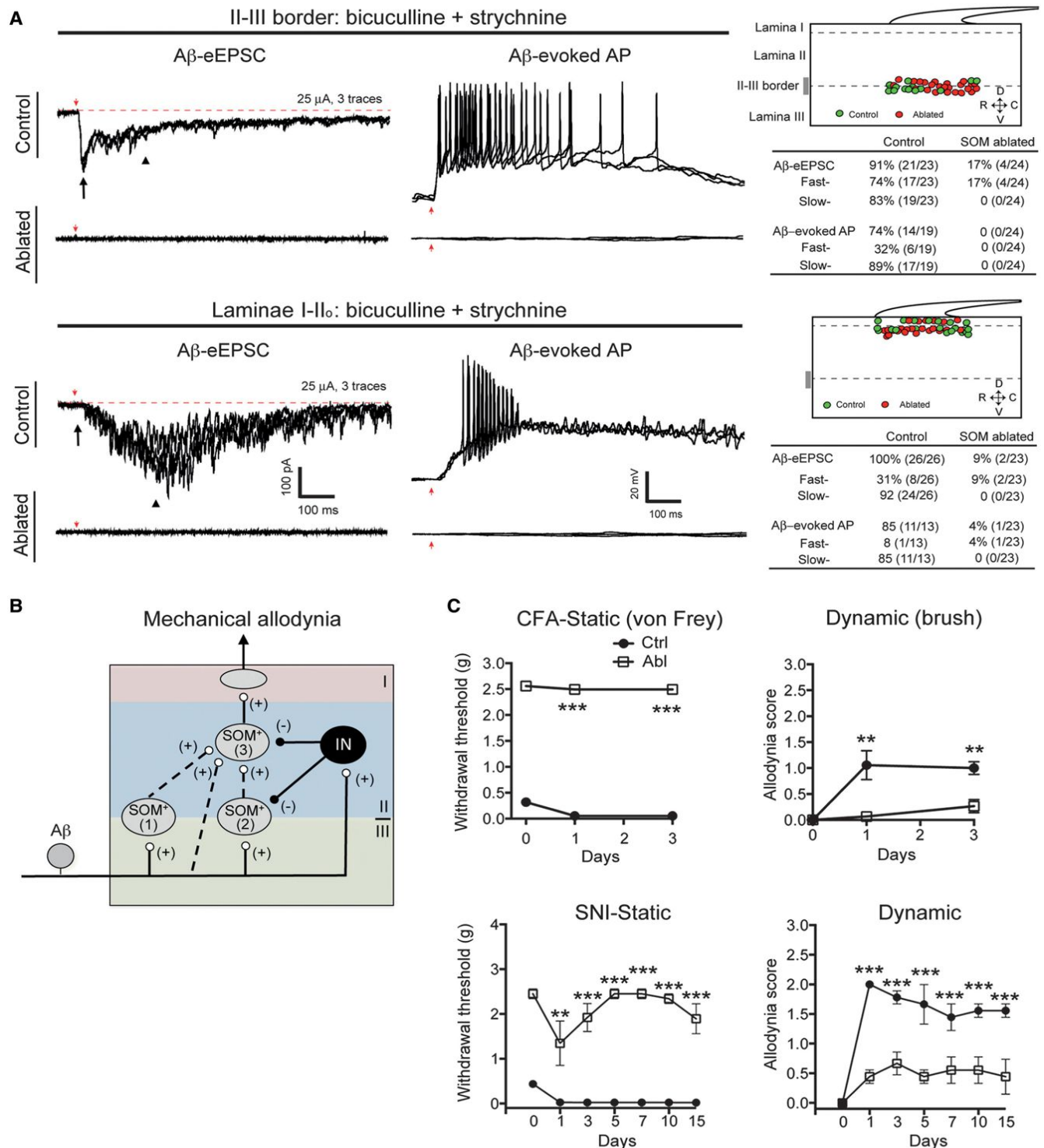


Figure 4. Loss of Aβ Inputs onto Lamina I/II Neurons and Mechanical Allodynia in SOM Abl Mice

(A) Aβ-evoked EPSCs/APs in spinal neurons from control and SOM Abl mice. Left: typical traces; right: the positions of recorded neurons and summary. Red arrows indicate stimulation artifacts. Black arrows and arrowheads indicate fast and slow eEPSCs, respectively.

(B) Schematic showing SOM neurons linking Aβ fibers to lamina I neurons, which is gated by inhibitory neurons.

(C) Loss of static (von Frey assay) and dynamic (brush assay) mechanical allodynia following peripheral inflammation and nerve injury in SOM Abl mice ("Abl," open rectangle) in comparison with control ("Ctrl," solid circles) ($n = 6-7$ in each group; $p < 0.001$, one-way ANOVA with Newman-Keuls post hoc analysis). Data are represented as mean \pm SEM. See also Figure S2.

(Figure 5B) and the hindbrain Sp5 nucleus, but not in other brain areas, and the ablation did not affect afferent projections (Figure S7B). However, DTX treatment did not ablate Dyn-Tomato⁺ VGLUT2⁺ glutamatergic neurons (Figure 5B), suggesting that these excitatory neurons might originate from Lbx1-negative spinal neurons (Gross et al., 2002; Müller et al., 2002).

Spontaneous Development of Mechanical Allodynia in Dyn Abl Mice

We next performed behavioral analyses in Dyn Abl mice, using littermates as controls. We found that Dyn Abl and control mice did not exhibit differences in locomotor coordination (data not shown), or in responses to heat or cold stimuli (Figures S7D–S7G). Furthermore, ablation of Dyn-Tomato⁺ neurons in adult mice did not change itch sensitivity (for discussion, see Figures S7H–S7M). Strikingly, Dyn Abl mice showed spontaneous development of both static and dynamic mechanical allodynia (Figure 5C). Moreover, the values of allodynia cannot be further increased by inflammation or nerve injury (Figure 5D and 5E). In contrast, no allodynia developed upon ablation of ChAT (Figure S6) or NPY (described elsewhere) lineages of inhibitory neurons. Thus, Dyn lineage neurons are uniquely required to gate mechanical pain.

Dyn-Tomato⁺ Neurons in Lamina II_o Receive A β Inputs with AP Firing

We next examined afferent inputs onto Dyn-Tomato⁺ neurons. In total, 103 Dyn-Tomato⁺ neurons from 16 mice at P24–P30 were recorded. At A β stimulation intensity range, eEPSCs can be detected in the vast majority of Dyn-Tomato⁺ neurons at -70 mV (Figure 6A). However, the strength of A β input is quite different in different laminae. In II_i and at the II–III border, most Dyn-Tomato⁺ neurons (22/24) receive A β input with small eEPSCs and feed-forward inhibition, and only a few of them (2/24) produce A β -evoked APs. In contrast, Dyn-Tomato⁺ neurons in laminae I and II_o receive monosynaptic or polysynaptic A β input with less feed-forward inhibition, and more than half of them produce A β -evoked APs (Figure 6A). Dorsally located Dyn-Tomato⁺ neurons include vertical cells that send dendrites all the way to laminae III–IV (Figure 6B), thereby forming an anatomical basis for receiving direct A β inputs.

Dyn-Tomato⁺ Neuron Ablation Leads to A β -Evoked AP Firing in Superficial Dorsal Horn Neurons

We next recorded from randomly picked neurons located at different laminae under normal ACSF recording conditions in control and Dyn Abl mice. At the II–III border, neurons receiving A β inputs are increased from 66% (25/38) in control mice to 94% (32/34) in Dyn Abl mice (Chi-square test, $p < 0.01$; Figure 7A), and neurons generating A β -evoked APs are increased by 35%, from 5% (2/38) in control mice to 40% (12/30) in Dyn Abl mice (Chi-square test, $p < 0.001$; Figure 7A). Note that A β stimulation evoked fast EPSCs in all responsive neurons at the II–III border in both control and ablation mice, but slow EPSCs only in a subset of neurons in Abl mice (15%; 5/34). Thus, Dyn neurons are required to gate A β inputs onto a portion of II–III border neurons.

In laminae I and II_o, neurons receiving A β inputs are increased from 36% (32/89) in control mice to 83% (43/52) in

Dyn Abl mice (Chi-square test, $p < 0.001$; Figure 7A). Moreover, A β stimulation only generated fast APs in 7% (5/69) of neurons in control mice but can generate fast and/or slow APs in 76% (34/45) in Dyn Abl mice (Chi-square test, $p < 0.001$; Figure 7A). More surprisingly, 31% (16/52) of I–II_o neurons received monosynaptic A β inputs with AP firing, which was rarely observed in control mice (1%; 1/69; Chi-square test, $p < 0.001$). It should be noted that for control mice recorded in the presence of bicuculline/strychnine, A β stimulation mainly evoked slow, but not fast, AP firing in I–II_o neurons (see above, Figure 4). Thus, Dyn lineage inhibitory neurons provide two gating mechanisms for I/II_o neurons: (1) a bicuculline/strychnine-sensitive one that prevents slow A β -evoked AP firing, and (2) a bicuculline/strychnine-insensitive one that prevents fast A β -evoked AP firing.

Low Threshold Mechanical Stimuli Activate SOM⁺ Neurons in Dyn Abl Mice

We next tested whether low-threshold mechanical force can activate SOM⁺ pain transmission neurons upon ablation of Dyn-Tomato⁺ inhibitory neurons. To do this, we brushed one side of the shaved back skin, and monitored the activation of spinal neurons by c-Fos induction. We found that this low-threshold brushing stimulus induced c-Fos in thoracic dorsal horn neurons of Dyn Abl mice, but rarely in control littermates (Figure 7B). Double immunostaining showed that $21.7\% \pm 3.4\%$ of these c-Fos⁺ neurons showed detectable expression of the SOM peptide (Figure 7C), while the few c-Fos⁺ neurons in control mice showed almost no SOM expression ($1.1\% \pm 1.1\%$). Thus, Dyn⁺ inhibitory neurons are required to prevent low-threshold mechanical stimuli from activating SOM⁺ pain transmission neurons (summarized in Figure 7D).

DISCUSSION

Our studies show that SOM lineage excitatory neurons, enriched in lamina II, are required to sense mechanical pain, but not thermal pain. SOM neurons are also part of polysynaptic circuits linking A β fibers to pain output neurons, and their ablation results in the loss of mechanical allodynia induced by inflammation or nerve lesions. Furthermore, we show that A β input onto superficial dorsal horn neurons is gated through feed-forward activation of the Dyn lineage inhibitory neurons.

Lamina Organization in Transmitting Mechanical Pain-Related Information

Dorsal horn neurons are organized into laminae (Rexed, 1952). Ascending projection neurons are enriched in laminae I and scattered throughout III–VI, whereas neurons in lamina II mainly belong to local interneurons (Braz et al., 2014; Todd, 2010; Willis et al., 2001). In a landmark study published in 1970, Christensen and Perl discovered that nociceptive neurons in lamina I either respond to noxious mechanical stimuli alone or are polymodal, responding to both noxious heat and mechanical stimuli (Christensen and Perl, 1970). Only a few spinothalamic projection neurons in lamina I respond selectively to noxious heat (Han et al., 1998). SOM-Tomato⁺ neurons are enriched in lamina II, with little overlap with NK1R⁺ lamina I ascending projection neurons. How

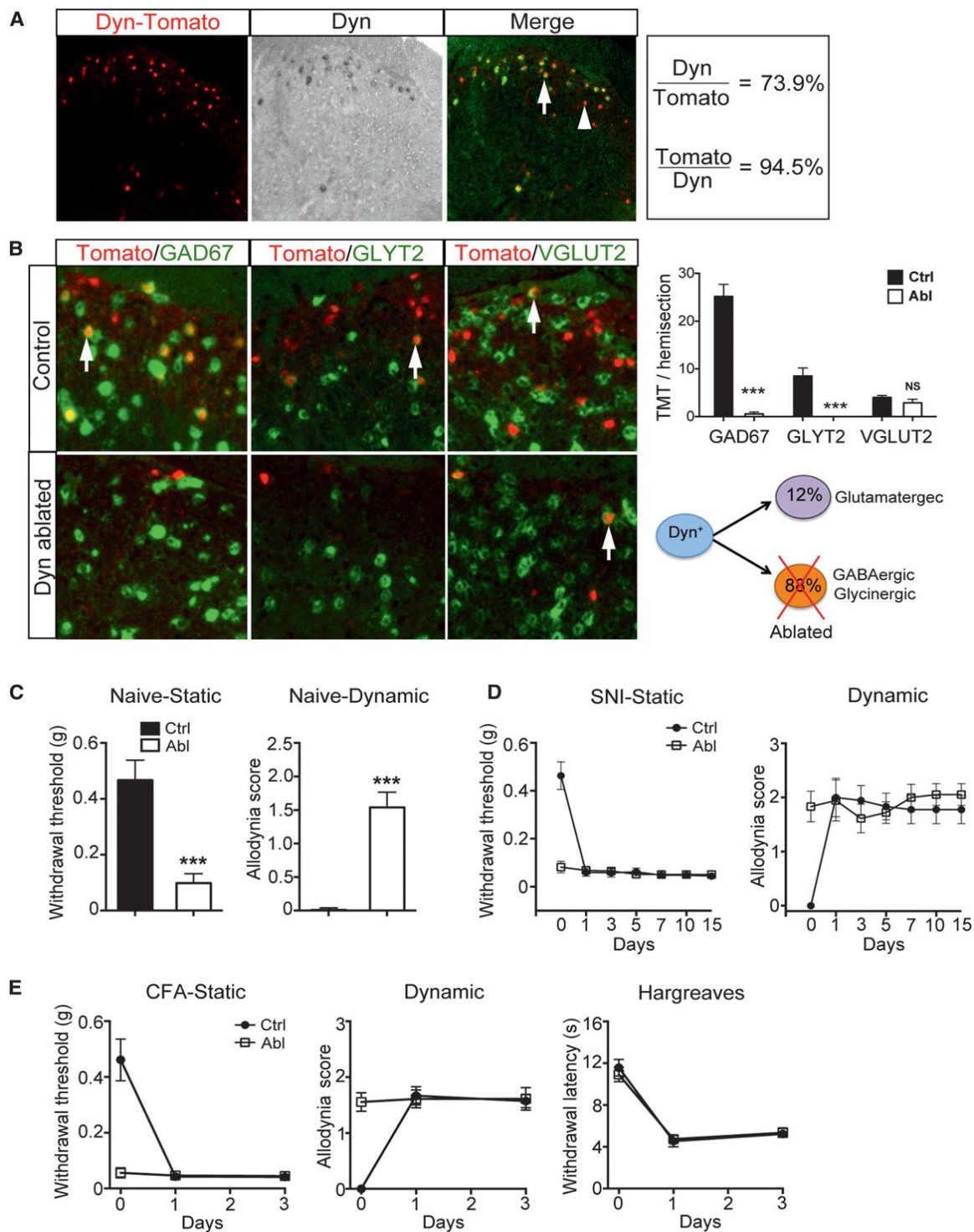


Figure 5. Spontaneous Development of Mechanical Allodynia in Dyn Abl Mice

(A) Double staining of Tomato and Dyn mRNA in the spinal cord of Dyn-Tomato mice.

(B) Double staining of Tomato and indicated mRNAs in Dyn-Tomato control mice and Dyn Abl mice. Right (upper): quantification analysis. Schematic in lower right showing selective ablation of inhibitory Dyn lineage neurons.

(C) Reduction of withdrawal threshold to static stimuli (von Frey assay) and increase in dynamic allodynia score (brush assay) in Dyn neuron-ablated ("Abl") mice (n = 15 in control ["Ctrl"] group, n = 10 in Abl group; ***p < 0.001, Student's unpaired t test).

(legend continued on next page)

then can mechanical pain be selectively lost following ablation of SOM neurons?

It should be noted that prior *in vivo* recordings have not been able to determine whether lamina I projection neurons receive mono- or polysynaptic inputs from primary afferents. In lamina II, neurons located in II_o and d-II_i predominantly receive nociceptive inputs, based on extracellular recording (Bennett et al., 1980; Cervero et al., 1979; Cervero et al., 1976; Christensen and Perl, 1970; Kumazawa and Perl, 1978; Light et al., 1979). Other studies indicate that vertical cells in II_o and d-II_i are the only output neurons that project their axons from lamina II to lamina I (Bennett et al., 1980; Gobel, 1978; Light et al., 1979; Lu and Perl, 2005; Molony et al., 1981; Price et al., 1979). We found that SOM neurons do include vertical cells (Figure S1E). Thus, mechanical nociceptors must transmit noxious mechanical information to lamina I projection neurons via lamina II SOM neurons (and/or those few SOM neurons scattered in laminae I and III-V) (Figure 2H), although it is not known whether SOM⁺ neurons connect with ascending projection neurons directly or indirectly. Furthermore, by comparing behavioral phenotypes of SOM versus Calb2 Abl mice, we reveal separate spinal neuronal populations transmitting light punctate versus intense noxious mechanical information (Figure S5L).

In contrast to the abolition of mechanical pain, SOM Abl mice show normal nocifensive responses to noxious heat and cold stimulation. Previous *in vivo* recordings showed that heat stimuli evoke firing in neurons located predominantly in lamina I/II_o and in lamina V, but only rarely in lamina II (Furie et al., 1999). The enrichment of SOM⁺ neurons in lamina II may explain why thermal pain is unaffected in SOM Abl mice, although our studies do not rule out a redundant role for SOM⁺ neurons in processing thermal information. The polymodal nature of lamina I and possibly lamina V neurons might be due to the convergence of direct inputs from heat fibers and indirect inputs from mechanosensitive nociceptors via lamina II SOM neurons. Thus, our studies gain insight into how different modalities of nociceptive information are transmitted through distinct dorsal horn laminae.

Identification of Spinal Circuits for Gate Control of Mechanical Pain

The gate control theory postulates that spinal pain transmission (T) neurons receive inputs from both nociceptors and A β mechanoreceptors, with A β inputs gated through feed-forward activation of spinal inhibitory neurons (INs). Our data suggest that the SOM lineage of excitatory neurons and the Dyn lineage of inhibitory neurons represent the T neurons and INs, respectively. The original gate theory designates T cells as the ascending projection neurons, but our studies show that T neurons can be lamina II interneurons. We found that lamina II SOM⁺ neurons receive monosynaptic inputs from mechanical nociceptors, as well as A β inputs with feed-forward inhibition. Strikingly, ablation of SOM neurons leads to a virtual loss of A β fiber inputs onto the

superficial dorsal horn. Consistently, chronic mechanical allodynia induced by inflammation or nerve lesions, which is partly caused by disinhibition that allows low-threshold mechanical stimuli to activate pain transmission neurons, is abolished in SOM Abl mice (Sandkühler, 2009; Woolf and Doubell, 1994; Zeilhofer et al., 2012). Several lines of evidence support the Dyn lineage neurons functioning as IN-type inhibitory neurons. First, A β stimulation is able to evoke AP firing in a subset of Dyn neurons. Second, ablation of Dyn neurons leads to A β -evoked AP firing in most lamina I/II neurons, leading to spontaneous development of mechanical allodynia. It should be noted that the Dyn peptide has both pronociceptive and antinociceptive roles (Lai et al., 2006). Our data suggest that the net output of the Dyn lineage neurons is, however, inhibitory. The induction of c-Fos in SOM⁺ neurons by skin brushing in Dyn Abl mice, but not in control mice, suggests that Dyn inhibitory neurons normally acts to prevent low-threshold mechanical stimuli from activating SOM⁺ pain transmission neurons.

Dyn neurons gate two polysynaptic excitatory circuits linking A β fibers to lamina I pain output neurons via lamina II SOM⁺ neurons. In pathway A ("A" in Figure 7D), Lu and others showed that A β fibers form monosynaptic connection to PKC γ ⁺ excitatory neurons at the II-III border, which are in turn connected to lamina II transient central and vertical cells, and finally to lamina I projection neurons (Lu et al., 2013). These PKC γ ⁺ neurons at the II-III border likely represent type 2 SOM-Tomato⁺ neurons described in Figure 3 since both types of neurons receive A β inputs that are gated through bicuculline/strychnine-sensitive feed-forward inhibition (Lu et al., 2013) and such A β inputs are completely lost in SOM Abl mice. Lu et al. further reported a ventrally located glycinergic inhibitory gate at the II-III border (Lu et al., 2013). Dyn⁺ neurons contribute to this ventral gate since 35% of II-III border neurons gain the ability to generate A β -evoked AP output in Dyn Abl mice. The degree of gate opening (35%) is, however, less than the 69% caused by bicuculline/strychnine treatment in control mice, indicating the existence of Dyn-independent ventral gates.

In pathway B ("B" in Figure 7D), A β fibers may provide direct inputs onto vertical cells located at II_o. Prior studies showed that II_o vertical cells send their dendrites ventrally, reaching the II-III border or even laminae III and IV (Bennett et al., 1980; Gobel, 1978; Light et al., 1979; Lu and Perl, 2005; Molony et al., 1981; Price et al., 1979). The presence of numerous spines in distal dendrites (Gobel, 1978) indicates that these neurons receive synaptic inputs from a region enriched with A β terminals, as confirmed by laser scanning photostimulation studies (Kato et al., 2009; Kosugi et al., 2013). This direct pathway likely depends on type 3 SOM neurons since these neurons include vertical cells (Figure S1E) and receive fast A β -evoked EPSCs (Figure 3B). However, to prevent pain from being evoked by innocuous mechanical stimuli, these vertical cells in II_o have to be gated, and we found that Dyn neurons are again involved. Indeed, to our surprise, Dyn neurons that receive A β input with

(D and E) After spared nerve injury (SNI) (D) or peripheral inflammation by CFA treatment (E), control mice show a reduction in withdrawal thresholds to static stimuli by von Frey assay and an increase of dynamic allodynia score by the brush assay ($n = 7$, $p < 0.001$, one-way ANOVA with Newman-Keuls post hoc analysis). No difference before and after nerve injury or inflammation in Abl mice ($n = 6$; $p > 0.05$, one-way ANOVA with Newman-Keuls post hoc analysis). Data are represented as mean \pm SEM. See also Figures S6 and S7.

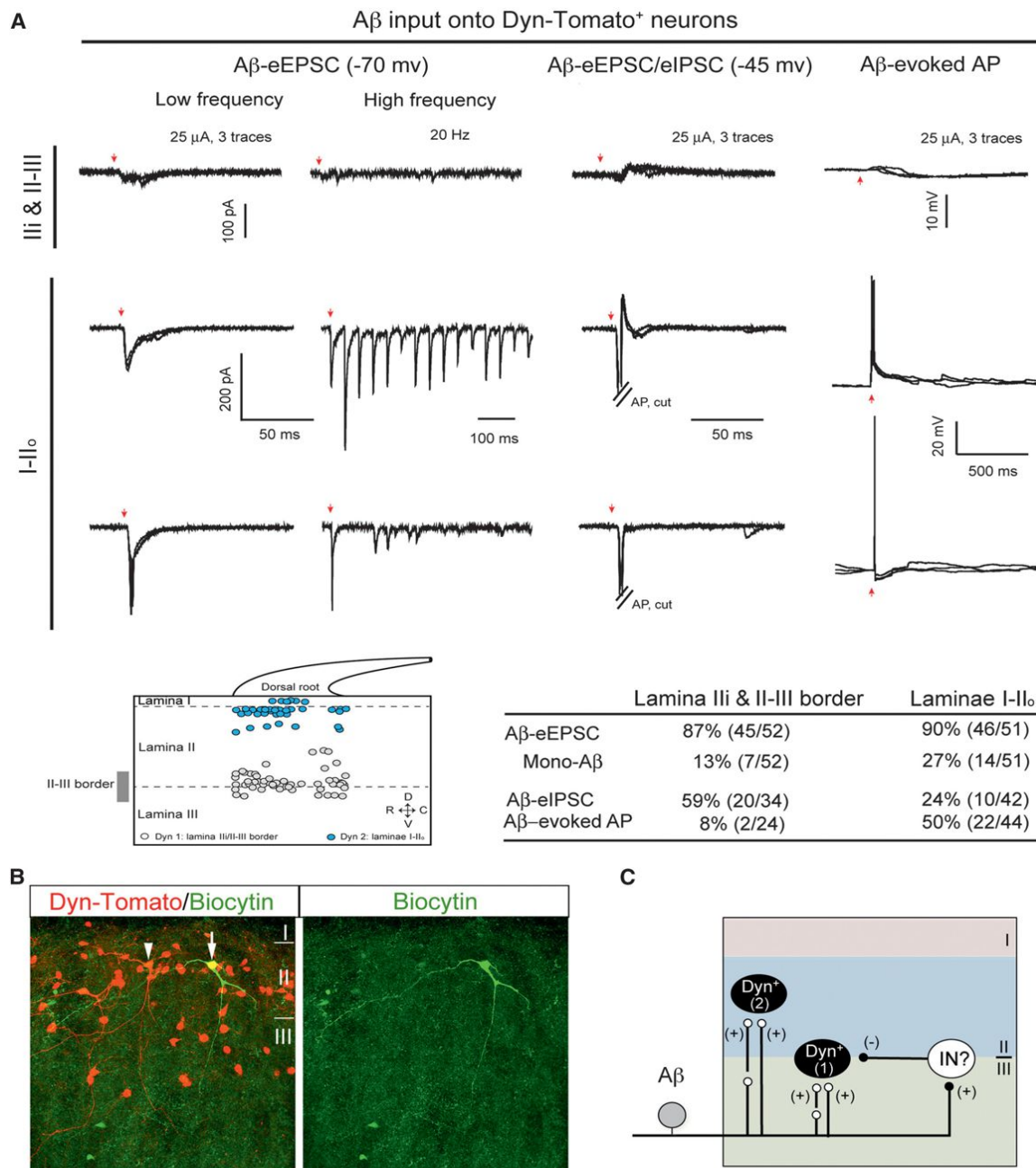


Figure 6. A β Input onto Dyn-Tomato⁺ Neurons

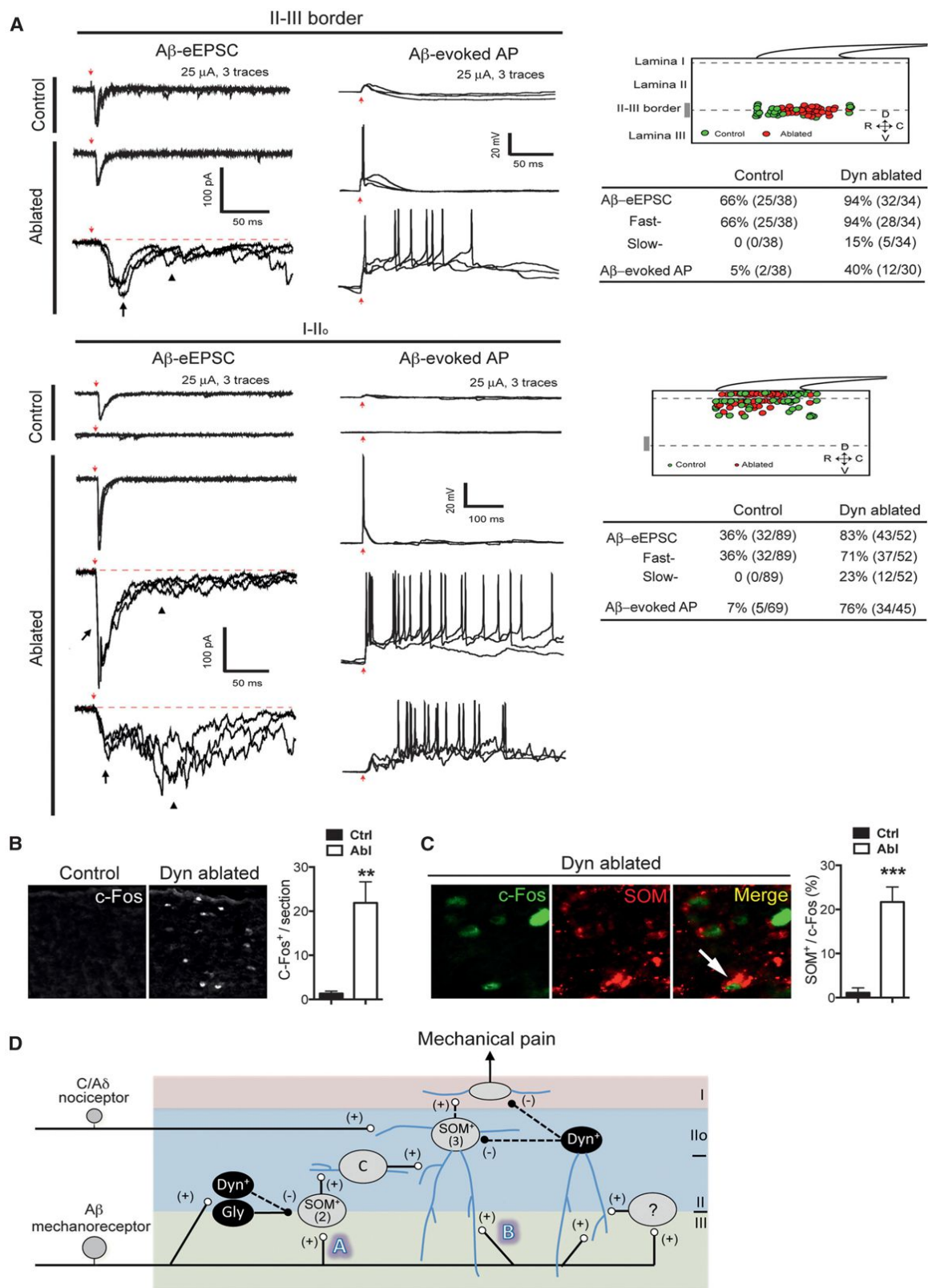
(A) Aβ input onto Dyn-Toamto⁺ neurons. Upper lanes show typical traces. Middle showing the relative positions and summary of recorded Dyn-Toamto⁺ neurons from 13 mice. Red arrows indicate stimulation artifacts.

(B) Biocytin labeling showing vertical dendritic arborization of a biocytin-injected Dyn-Tomato⁺ neuron (arrow) and an uninjected Dyn-Tomato⁺ neuron (arrowhead) in lamina II_o.

(C) Schematic showing Dyn-Tomato⁺ neurons in lamina II_i and at II-III border ("1") that receive A β input with strong feed-forward inhibition, and in laminae I-II_o ("2") that receive A β input with AP output. Not shown are small subsets of types 1 and 2 cells located in I/II_o and at II-III border, respectively.

AP output are mainly located in laminae I and II_o. Moreover, dorsally located Dyn neurons include vertical cells with ventrally projected dendrites, thereby placing them in a perfect position to receive direct Aβ inputs, either directly or indirectly (i.e., via type 1 SOM⁺ neurons), and feed-forwardly inhibit nearby vertical

excitatory neurons (Figure 7D). Indeed, this direct A β -evoked pathway B becomes open in Dyn Abl mice, as indicated by 31% of lamina II α neurons that gain the ability to receive fast monosynaptic A β inputs with AP firing. Interestingly, this Dyn-dependent dorsal gate is mediated through a mechanism



(legend on next page)

that is insensitive to bicuculline/strychnine, thereby distinguishing it from the ventral gate that is sensitive to bicuculline/strychnine. Direct A β -evoked inhibitory inputs onto II $_o$ nociceptive neurons can explain why A β stimulation has analgesic effects (Bini et al., 1984; Head, 1905; Salter and Henry, 1990; Wall and Sweet, 1967).

CONCLUSIONS

This study gives us an opportunity to remember the wisdom articulated in the 1960s and 1970s by two late sensory titans: Patrick Wall and Edward Perl. The requirement of SOM neurons for sensing mechanical pain, but not touch or temperature, supports the existence of specific pain-related circuits argued by Perl. The enrichment of SOM neurons in the substantia gelatinosa (Rexed's lamina II) also suggests a critical role of this lamina in processing mechanical pain. SOM neurons are heterogeneous and further studies are warranted to determine if they transmit other modalities, such as mechanical itch. Meanwhile, the finding that SOM neurons receive A β inputs with feed-forward inhibition via Dyn inhibitory neurons supports the core argument of the gate theory proposed by Wall (and Melzack). Thus, pain is encoded through a hybrid mechanism that combines Perl's specificity and Wall's pattern theories, a mechanism recently referred to as the population coding theory (Ma, 2010, 2012; Prescott et al., 2014). Clinically, mechanical pain treatment represents a big challenge (Lolignier et al., 2014). Our study suggests that drugs targeted at reducing excitatory output from SOM neurons or enhancing inhibitory output from Dyn neurons could be ideally used to attenuate mechanical allodynia, without affecting the senses of temperature and touch that are vital for daily life.

EXPERIMENTAL PROCEDURES

Genetic Marking and Ablation of Spinal Neurons

The SOM, Tac2, Calb2, Dyn, ChAT and NPY lineage neurons in the dorsal spinal cord were labeled by crossing various *Cre* lines with the *tdTomato* reporter line. These *Cre* lines were then crossed with *Tau-DTR* and *Lbx1-Flpo* mice to drive DTR expression selectively in specific spinal lineage neurons. DTR-expressing neurons were ablated upon intraperitoneal injection with diphtheria toxin (DTX) at day 1 and day 4. Details of mouse lines and intersectional ablation could be found in the Extended Experimental Procedures.

In Situ Hybridization, Immunohistochemistry

ISH and IHC (Liu et al., 2010) were performed using standard methods (see Extended Experimental Procedures).

Behavioral Testing

Surgery and behavior testing were performed as previously described (Knowlton et al., 2013; Liu et al., 2010). Sensorimotor coordination was measured by rotarod, innocuous touch sensations were measured using sticky tape and brushing assays, thermal sensations were measured by the Hargreaves, hot plate, cold plate, and acetone evaporation assays, and mechanical pain were measured by von Frey, pinprick, and pinch assays. Static allodynia was measured using von Frey assay, and dynamic allodynia was measured using the scoring system developed by Dr. Enrique José Cobos (see Extended Experimental Procedures for details).

Spinal Cord Slice Preparation, Patch Clamp Recording, and Biocytin Labeling

The lumbar spinal cord of mice (P23-P30) was removed and then sagittal spinal cord slices (350–500 μ m) with dorsal roots (8–18 mm) attached were cut. Whole-cell patch-clamp recordings were performed. To reveal neuron morphology, biocytin was filled in the targeted neuron after a minimum of 20 min in the whole-cell, tight-seal patch-clamp configuration (see Extended Experimental Procedures for details).

Statistical Analysis

Results are expressed as mean \pm SEM. The $p < 0.05$ was accepted as statistically different (see Extended Experimental Procedures for details).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures and seven figures and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2014.11.003>.

AUTHOR CONTRIBUTION

B.D., S.B., W.K., and X.R. performed histochemical and behavioral analyses; L.C. performed electrophysiological recording; O.B., C.P., L.G.-C., M.K., T.V., S.R., and B.B.L. provided unpublished mouse lines; Q.M., M.G. and Y.W. supervised the whole study; Q.M., B.D., L.C., S.B. and M.G. wrote the manuscript.

ACKNOWLEDGMENTS

We thank Dr. Z. Josh Huang and the Jackson laboratory for the *SOM-IRES-Cre* and *Calb2-IRES-Cre* mice, GENSAT and MMRRRC at University of California, Davis, for the *NPY-Cre* mice, the Allen Brain Institute, and the Jackson Laboratory for the *Rosa26^{LSL-tdTomato}* mice, and Dr. Susan Dymecki for the

Figure 7. Dyn Neurons Gate Mechanical Pain

(A) A β -evoked EPSCs/APs in the spinal dorsal horn of control and Dyn Abl mice. Left: typical traces. Right: positions of recorded neurons and summary. Twenty-seven control mice and 11 ablated mice were used. Red arrows indicate stimulation artifacts. Black arrows indicate fast eEPSCs. Arrowheads indicated slow eEPSCs.

(B) Brush-evoked c-Fos induction in the dorsal spinal cord of Dyn Abl and control mice [n = 12 sections in control ("Ctrl") group, n = 9 sections in Abl group, 3 mice in each group; **p < 0.01, Student's unpaired t test].

(C) Double immunostaining of c-Fos with SOM (arrow) following back brush stimuli in Dyn Abl and control mice (n = 12 thoracic spinal sections in Ctrl and Abl groups, 4 mice in each group; ***p < 0.001, Student's unpaired t test).

(D) Schematic showing circuitry processing mechanical pain-related information. Vertical neurons in lamina II $_o$, belonging to type 3 SOM⁺ neurons ["(3)"], receive inputs from C/A δ mechanical nociceptors, and also from A β mechanoreceptors through two pathways: indirect ("A") and direct ("B"). Pathway "A" is transmitted through type 2 SOM⁺ ["(2)"] neurons at the II-III border, via transient-central ("C") cells and vertical cells in lamina II $_o$, and finally to lamina I projection neurons, although it is not known if the connection from vertical cells to projection neurons is direct or indirect. Type 2 SOM⁺ neurons may include PKC γ ⁺ neurons. Pathway "A" is partly gated by Dyn⁺ neurons. Pathway "B" is indicated by direct A β inputs onto lamina II $_o$ neurons, and is gated by dorsally located Dyn⁺ neurons that receive A β inputs with AP output, either directly or via type 1 SOM⁺ neurons or unidentified interneurons ("?"). Dashed arrows indicate that SOM⁺ neurons might receive direct inhibitory inputs from Dyn⁺ neurons, but further studies are required to confirm this. Our data do not rule out that Dyn⁺ neurons might also directly gate lamina I projection neurons. For details, see Discussion. Data are represented as mean \pm SEM.

ROSA26^{CAG-FRT-STOP-FRT-GFP} mice. We thank Drs. Yan Lu, Clifford Woolf, and Fu-Chia Yang for critical comments on the manuscript; Dr Yan Lu for his advice on spinal cord slice recording; and Dr. Enrique José Cobos for providing the scoring system in measuring dynamic allodynia. The Ma lab was supported by NIH grants (R01NS086372, NS047710, and P01 NS072040), the Goulding lab is supported by NIH grants (R01NS086372, R01 NS 080586, and P01 NS072031), the Ross lab is supported by NIH grants (R01 AR063772 and R21 AR064445), and the Lowell lab was supported by NIH grants (R01 DK075632, R37 DK053477, R01 DK071051, R01 DK089044, R01 DK096010, P30 DK057521, and P30 DK046200). M.K. was supported by F32 DK089710. L.C. and Y.W. were supported by grants from the National Natural Science Foundation of China (81171224 and 81100815).

Received: August 19, 2014

Revised: October 7, 2014

Accepted: October 31, 2014

Published: November 20, 2014

REFERENCES

- Abaira, V.E., and Ginty, D.D. (2013). The sensory neurons of touch. *Neuron* 79, 618–639.
- Baba, H., Ji, R.R., Kohno, T., Moore, K.A., Ataka, T., Wakai, A., Okamoto, M., and Woolf, C.J. (2003). Removal of GABAergic inhibition facilitates polysynaptic A fiber-mediated excitatory transmission to the superficial spinal dorsal horn. *Mol. Cell. Neurosci.* 24, 818–830.
- Bennett, G.J., Abdelmoumene, M., Hayashi, H., and Dubner, R. (1980). Physiology and morphology of substantia gelatinosa neurons intracellularly stained with horseradish peroxidase. *J. Comp. Neurol.* 194, 809–827.
- Bessou, P., and Perl, E.R. (1969). Response of cutaneous sensory units with unmyelinated fibers to noxious stimuli. *J. Neurophysiol.* 32, 1025–1043.
- Bini, G., Cruccu, G., Hagbarth, K.E., Schady, W., and Torebjörk, E. (1984). Analgesic effect of vibration and cooling on pain induced by intraneural electrical stimulation. *Pain* 18, 239–248.
- Braz, J., Solorzano, C., Wang, X., and Basbaum, A.I. (2014). Transmitting pain and itch messages: a contemporary view of the spinal cord circuits that generate gate control. *Neuron* 82, 522–536.
- Bruno, R.M. (2011). Synchrony in sensation. *Curr. Opin. Neurobiol.* 21, 701–708.
- Burgess, P.R., and Perl, E.R. (1967). Myelinated afferent fibres responding specifically to noxious stimulation of the skin. *J. Physiol.* 190, 541–562.
- Campbell, J.N., and Meyer, R.A. (2006). Mechanisms of neuropathic pain. *Neuron* 52, 77–92.
- Campbell, J.N., Raja, S.N., Meyer, R.A., and Mackinnon, S.E. (1988). Myelinated afferents signal the hyperalgesia associated with nerve injury. *Pain* 32, 89–94.
- Carstens, E.E., Carstens, M.I., Simons, C.T., and Jinks, S.L. (2010). Dorsal horn neurons expressing NK-1 receptors mediate scratching in rats. *Neuroreport* 21, 303–308.
- Cervero, F., Iggo, A., and Ogawa, H. (1976). Nociceptor-driven dorsal horn neurones in the lumbar spinal cord of the cat. *Pain* 2, 5–24.
- Cervero, F., Iggo, A., and Molony, V. (1979). An electrophysiological study of neurones in the Substantia Gelatinosa Rolandi of the cat's spinal cord. *Q. J. Exp. Physiol. Cogn. Med. Sci.* 64, 297–314.
- Chaplan, S.R., Bach, F.W., Pogrel, J.W., Chung, J.M., and Yaksh, T.L. (1994). Quantitative assessment of tactile allodynia in the rat paw. *J. Neurosci. Methods* 53, 55–63.
- Christensen, B.N., and Perl, E.R. (1970). Spinal neurons specifically excited by noxious or thermal stimuli: marginal zone of the dorsal horn. *J. Neurophysiol.* 33, 293–307.
- Decosterd, I., and Woolf, C.J. (2000). Spared nerve injury: an animal model of persistent peripheral neuropathic pain. *Pain* 87, 149–158.
- Dymecki, S.M., and Kim, J.C. (2007). Molecular neuroanatomy's "Three Gs": a primer. *Neuron* 54, 17–34.
- Freneau, R.T.J., Jr., Voglmaier, S., Seal, R.P., and Edwards, R.H. (2004). VGLUTs define subsets of excitatory neurons and suggest novel roles for glutamate. *Trends Neurosci.* 27, 98–103.
- Furue, H., Narikawa, K., Kumamoto, E., and Yoshimura, M. (1999). Responsiveness of rat substantia gelatinosa neurones to mechanical but not thermal stimuli revealed by in vivo patch-clamp recording. *J. Physiol.* 521, 529–535.
- Gobel, S. (1978). Golgi studies of the neurons in layer II of the dorsal horn of the medulla (trigeminal nucleus caudalis). *J. Comp. Neurol.* 180, 395–413.
- Gross, M.K., Dottori, M., and Goulding, M. (2002). Lbx1 specifies somatosensory association interneurons in the dorsal spinal cord. *Neuron* 34, 535–549.
- Han, Z.S., Zhang, E.T., and Craig, A.D. (1998). Nociceptive and thermoreceptive lamina I neurons are anatomically distinct. *Nat. Neurosci.* 1, 218–225.
- Head, H. (1905). The afferent nervous system from a new aspect. *Brain* 28, 100–115.
- Kato, G., Kawasaki, Y., Koga, K., Uta, D., Kosugi, M., Yasaka, T., Yoshimura, M., Ji, R.R., and Strassman, A.M. (2009). Organization of intralaminar and translaminal neuronal connectivity in the superficial spinal dorsal horn. *J. Neurosci.* 29, 5088–5099.
- Knowlton, W.M., Palkar, R., Lippoldt, E.K., McCoy, D.D., Baluch, F., Chen, J., and McKemy, D.D. (2013). A sensory-labeled line for cold: TRPM8-expressing sensory neurons define the cellular basis for cold, cold pain, and cooling-mediated analgesia. *J. Neurosci.* 33, 2837–2848.
- Koltzenburg, M., Lundberg, L.E., and Torebjörk, H.E. (1992). Dynamic and static components of mechanical hyperalgesia in human hairy skin. *Pain* 51, 207–219.
- Kosugi, M., Kato, G., Lukashov, S., Pendse, G., Puskar, Z., Kozsurek, M., and Strassman, A.M. (2013). Subpopulation-specific patterns of intrinsic connectivity in mouse superficial dorsal horn as revealed by laser scanning photostimulation. *J. Physiol.* 591, 1935–1949.
- Krashes, M.J., Shah, B.P., Madara, J.C., Olson, D.P., Strohlic, D.E., Garfield, A.S., Vong, L., Pei, H., Watabe-Uchida, M., Uchida, N., et al. (2014). An excitatory paraventricular nucleus to AgRP neuron circuit that drives hunger. *Nature* 507, 238–242.
- Kumazawa, T., and Perl, E.R. (1978). Excitation of marginal and substantia gelatinosa neurons in the primate spinal cord: indications of their place in dorsal horn functional organization. *J. Comp. Neurol.* 177, 417–434.
- Lai, J., Luo, M.C., Chen, Q., Ma, S., Gardell, L.R., Ossipov, M.H., and Porreca, F. (2006). Dynorphin A activates bradykinin receptors to maintain neuropathic pain. *Nat. Neurosci.* 9, 1534–1540.
- Light, A.R., Trevino, D.L., and Perl, E.R. (1979). Morphological features of functionally defined neurons in the marginal zone and substantia gelatinosa of the spinal dorsal horn. *J. Comp. Neurol.* 186, 151–171.
- Liu, Y., Abdel Samad, O., Zhang, L., Duan, B., Tong, Q., Lopes, C., Ji, R.R., Lowell, B.B., and Ma, Q. (2010). VGLUT2-dependent glutamate release from nociceptors is required to sense pain and suppress itch. *Neuron* 68, 543–556.
- Lolignier, S., Eijkelkamp, N., and Wood, J.N. (2014). Mechanical allodynia. *Pflugers Arch.* Published online May 22, 2014. <http://dx.doi.org/10.1007/s00424-014-1532-0>.
- Lu, Y., and Perl, E.R. (2005). Modular organization of excitatory circuits between neurons of the spinal superficial dorsal horn (laminae I and II). *J. Neurosci.* 25, 3900–3907.
- Lu, Y., Dong, H., Gao, Y., Gong, Y., Ren, Y., Gu, N., Zhou, S., Xia, N., Sun, Y.Y., Ji, R.R., and Xiong, L. (2013). A feed-forward spinal cord glycinergic neural circuit gates mechanical allodynia. *J. Clin. Invest.* 123, 4050–4062.
- Ma, Q. (2010). Labeled lines meet and talk: population coding of somatic sensations. *J. Clin. Invest.* 120, 3773–3778.
- Ma, Q. (2012). Population coding of somatic sensations. *Neurosci. Bull.* 28, 91–99.
- Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., et al. (2010). A robust

- and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* 13, 133–140.
- Mantyh, P.W., Rogers, S.D., Honore, P., Allen, B.J., Ghilardi, J.R., Li, J., Daughters, R.S., Lappi, D.A., Wiley, R.G., and Simone, D.A. (1997). Inhibition of hyperalgesia by ablation of lamina I spinal neurons expressing the substance P receptor. *Science* 278, 275–279.
- Mar, L., Yang, F.C., and Ma, Q. (2012). Genetic marking and characterization of Tac2-expressing neurons in the central and peripheral nervous system. *Mol. Brain* 5, 3.
- Melzack, R., and Wall, P.D. (1965). Pain mechanisms: a new theory. *Science* 150, 971–979.
- Melzack, R., and Wall, P.D. (1982). *The challenge of pain* (New York: Basic Books).
- Mendell, L.M. (2014). Constructing and deconstructing the gate theory of pain. *Pain* 155, 210–216.
- Miraucourt, L.S., Dallel, R., and Voisin, D.L. (2007). Glycine inhibitory dysfunction turns touch into pain through PKCgamma interneurons. *PLoS ONE* 2, e1116.
- Mishra, S.K., and Hoon, M.A. (2013). The cells and circuitry for itch responses in mice. *Science* 340, 968–971.
- Molony, V., Steedman, W.M., Cervero, F., and Iggo, A. (1981). Intracellular marking of identified neurones in the superficial dorsal horn of the cat spinal cord. *Q. J. Exp. Physiol.* 66, 211–223.
- Müller, T., Brohmann, H., Pierani, A., Heppenstall, P.A., Lewin, G.R., Jessell, T.M., and Birchmeier, C. (2002). The homeodomain factor *Ibx1* distinguishes two major programs of neuronal differentiation in the dorsal spinal cord. *Neuron* 34, 551–562.
- Noordenbos, W. (1987). Some historical aspects. *Pain* 29, 141–150.
- Prescott, S.A., Ma, Q., and De Koninck, Y. (2014). Normal and abnormal coding of somatosensory stimuli causing pain. *Nat. Neurosci.* 17, 183–191.
- Price, D.D., Hayashi, H., Dubner, R., and Ruda, M.A. (1979). Functional relationships between neurons of marginal and substantia gelatinosa layers of primate dorsal horn. *J. Neurophysiol.* 42, 1590–1608.
- Price, T.J., Cervero, F., Gold, M.S., Hammond, D.L., and Prescott, S.A. (2009). Chloride regulation in the pain pathway. *Brain Res. Brain Res. Rev.* 60, 149–170.
- Rexed, B. (1952). The cytoarchitectonic organization of the spinal cord in the cat. *J. Comp. Neurol.* 96, 414–495.
- Ribeiro-da-Silva, A., and De Koninck, Y. (2008). Morphological and neurochemical organization of the spinal dorsal horn. In *The Senses: A Comprehensive Reference, Volume 5*, Bushnell B.M.C. and Basbaum A.I., eds. (Amsterdam: Elsevier), pp. 279–310.
- Rossi, J., Balthasar, N., Olson, D., Scott, M., Berglund, E., Lee, C.E., Choi, M.J., Lauzon, D., Lowell, B.B., and Elmquist, J.K. (2011). Melanocortin-4 receptors expressed by cholinergic neurons regulate energy balance and glucose homeostasis. *Cell Metab.* 13, 195–204.
- Saito, M., Iwakaki, T., Taya, C., Yonekawa, H., Noda, M., Inui, Y., Mekada, E., Kimata, Y., Tsuru, A., and Kohno, K. (2001). Diphtheria toxin receptor-mediated conditional and targeted cell ablation in transgenic mice. *Nat. Biotechnol.* 19, 746–750.
- Salter, M.W., and Henry, J.L. (1990). Differential responses of nociceptive vs. non-nociceptive spinal dorsal horn neurones to cutaneously applied vibration in the cat. *Pain* 40, 311–322.
- Sandkühler, J. (2009). Models and mechanisms of hyperalgesia and allodynia. *Physiol. Rev.* 89, 707–758.
- Sardell, T.C., Polgár, E., Garzillo, F., Furuta, T., Kaneko, T., Watanabe, M., and Todd, A.J. (2011). Dynorphin is expressed primarily by GABAergic neurons that contain galanin in the rat dorsal horn. *Mol. Pain* 7, 76.
- Sun, Y.G., Zhao, Z.Q., Meng, X.L., Yin, J., Liu, X.Y., and Chen, Z.F. (2009). Cellular basis of itch sensation. *Science* 325, 1531–1534.
- Taniguchi, H., He, M., Wu, P., Kim, S., Paik, R., Sugino, K., Kvitsiani, D., Fu, Y., Lu, J., Lin, Y., et al. (2011). A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. *Neuron* 71, 995–1013.
- Todd, A.J. (2010). Neuronal circuitry for pain processing in the dorsal horn. *Nat. Rev. Neurosci.* 11, 823–836.
- Torsney, C., and MacDermott, A.B. (2006). Disinhibition opens the gate to pathological pain signaling in superficial neurokinin 1 receptor-expressing neurons in rat spinal cord. *J. Neurosci.* 26, 1833–1843.
- Wall, P.D. (1978). The gate control theory of pain mechanisms. A re-examination and re-statement. *Brain* 101, 1–18.
- Wall, P.D., and Sweet, W.H. (1967). Temporary abolition of pain in man. *Science* 155, 108–109.
- Wang, X., Zhang, J., Eberhart, D., Urban, R., Meda, K., Solorzano, C., Yamana, H., Rice, D., and Basbaum, A.I. (2013). Excitatory superficial dorsal horn interneurons are functionally heterogeneous and required for the full behavioral expression of pain and itch. *Neuron* 78, 312–324.
- Willis, W.D.J., Jr., Zhang, X., Honda, C.N., and Giesler, G.J.J., Jr. (2001). Projections from the marginal zone and deep dorsal horn to the ventrobasal nuclei of the primate thalamus. *Pain* 92, 267–276.
- Woolf, C.J., and Doubell, T.P. (1994). The pathophysiology of chronic pain—increased sensitivity to low threshold A beta-fibre inputs. *Curr. Opin. Neurobiol.* 4, 525–534.
- Xu, Y., Lopes, C., Wende, H., Guo, Z., Cheng, L., Birchmeier, C., and Ma, Q. (2013). Ontogeny of excitatory spinal neurons processing distinct somatic sensory modalities. *J. Neurosci.* 33, 14738–14748.
- Yasaka, T., Tiong, S.Y., Hughes, D.I., Riddell, J.S., and Todd, A.J. (2010). Populations of inhibitory and excitatory interneurons in lamina II of the adult rat spinal dorsal horn revealed by a combined electrophysiological and anatomical approach. *Pain* 151, 475–488.
- Yoshimura, M., and Nishi, S. (1995). Primary afferent-evoked glycine- and GABA-mediated IPSPs in substantia gelatinosa neurones in the rat spinal cord in vitro. *J. Physiol.* 482, 29–38.
- Zeilhofer, H.U., Wildner, H., and Yébenes, G.E. (2012). Fast synaptic inhibition in spinal sensory processing and pain control. *Physiol. Rev.* 92, 193–235.

A Constant Size Extension Drives Bacterial Cell Size Homeostasis

Manuel Campos,^{1,2,5} Ivan V. Surovtsev,^{1,2,5} Setsu Kato,^{3,5} Ahmad Paintdakhi,^{1,2} Bruno Beltran,^{1,2,3,6} Sarah E. Ebmeier,^{1,3} and Christine Jacobs-Wagner^{1,2,3,4,*}

¹Microbial Sciences Institute, Yale University, West Haven, CT 06516, USA

²Howard Hughes Medical Institute

³Department of Molecular, Cellular, and Developmental Biology

⁴Department of Microbial Pathogenesis, Yale Medical School

Yale University, New Haven, CT 06520, USA

⁵Co-first author

⁶Present address: Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803, USA

*Correspondence: christine.jacobs-wagner@yale.edu

<http://dx.doi.org/10.1016/j.cell.2014.11.022>

SUMMARY

Cell size control is an intrinsic feature of the cell cycle. In bacteria, cell growth and division are thought to be coupled through a cell size threshold. Here, we provide direct experimental evidence disproving the critical size paradigm. Instead, we show through single-cell microscopy and modeling that the evolutionarily distant bacteria *Escherichia coli* and *Caulobacter crescentus* achieve cell size homeostasis by growing, on average, the same amount between divisions, irrespective of cell length at birth. This simple mechanism provides a remarkably robust cell size control without the need of being precise, abating size deviations exponentially within a few generations. This size homeostasis mechanism is broadly applicable for symmetric and asymmetric divisions, as well as for different growth rates. Furthermore, our data suggest that constant size extension is implemented at or close to division. Altogether, our findings provide fundamentally distinct governing principles for cell size and cell-cycle control in bacteria.

INTRODUCTION

Cell size control, a universal property of all organisms, reflects the balance between growth and division. Mechanisms must be in place to ensure that cells narrowly distribute around a characteristic size for a given cell type, species, and growth condition. This is especially important for exponentially growing cells. Exponential growth implies that growth is proportional to cell size such that short cells grow slower than long cells in absolute growth rate. Thus, if no compensation occurs, any deviations from the mean size will increase cell size variability in the population at each generation. The very existence of a stable cell size distribution indicates the presence of intrinsic mechanisms that reduce cell size fluctuations.

Most cells—from bacteria to yeast to mammalian cells—are thought to regulate their size and cell cycle through critical size thresholds (Turner et al., 2012). In the critical size model, cells commit to division upon reaching a size threshold. Thus, all cells divide at about the same size whether they are born shorter or longer than the mean, compensating for their initial size deviation. The size threshold, or “sizer,” can be applied to a cell-cycle event other than division, with completion of this earlier event licensing cell division to occur after a constant amount of time, or “timer,” has elapsed. For example, the yeast *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* display a size threshold at the G1-S transition and mitosis, respectively (Fantès, 1977; Johnston et al., 1977; Sveiczer et al., 1996). In the bacterial field, a “sizer + timer” model gained momentum with seminal population studies in *Escherichia coli* and *Salmonella typhimurium*. A size threshold at the initiation of DNA replication was inferred from calculations showing that, on average, DNA replication initiates at a constant cell mass under different growth rate conditions (Donachie, 1968). Although disputed (Bates and Kleckner, 2005; Boye and Nordström, 2003; Wold et al., 1994), a coupling of cell division to DNA replication through a fixed timer was suggested from experiments showing that the timing between DNA replication and cell division remains constant across different growth rates (Cooper and Helmstetter, 1968; Schaechter et al., 1958; Schaechter et al., 1962). These findings observed at the population level were then assumed to be applicable to individual cells.

Cell size homeostasis could, at least in theory, be achieved through mechanisms that do not involve the licensing of division upon attainment of a certain size. These alternative mechanisms include a molecular clock, a simple timer, the addition of a constant cell volume, transition probability, or a concerted “sloppy” sizer and timer (Fantès and Nurse, 1981; Osella et al., 2014). For example, based on mathematical modeling, Voorn and Koppes first (Voorn and Koppes, 1998), and Amir later (Amir, 2014) argued that addition of a constant volume at each generation can describe the experimental shape of bacterial cell size distributions as well as population-derived bulk correlations (the positive correlation in size between mothers and daughters and the negative correlation between cell-cycle time and size at birth). However, these statistical features have alternative explanations

(Hosoda et al., 2011; Osella et al., 2014) and can be described by sizer-based homeostasis mechanisms (Koch and Schaechter, 1962; Koppes et al., 1980; Robert et al., 2014; Turner et al., 2012). The concept of sizer-based control has prevailed in the bacterial literature and, apart from the exception of *Mycobacterium* (Santi et al., 2013), still persists today as an underlying assumption in virtually all bacterial cell size and cell-cycle studies.

Importantly, beyond their associated caveats, all bacterial cell size homeostasis models—including the prevalent sizer-based models—lack direct experimental evidence. A direct examination of an intrinsic cell size mechanism requires the ability to track a large number of individual cells. Cell tracking must occur in the absence of environmentally induced cell size fluctuations. This is important because *E. coli* and other bacteria traditionally used for cell size studies change their average size in response to nutrient availability and cell density (Akerlund et al., 1995; Schaechter et al., 1958). Finally, cell size must be measured with high precision and at high temporal resolution, preferably over multiple cell cycles.

Another overlooked aspect of bacterial cell size homeostasis is the role of division site placement. Most bacterial cell size studies—whether experimental or theoretical—have focused on symmetrically dividing bacteria, even though asymmetric divisions are well represented in the bacterial world. For example, the large class of α -proteobacteria appears to be dominated by asymmetrically dividing bacterial species. The best-studied α -proteobacterium is *Caulobacter crescentus*, which divides asymmetrically to produce two daughter cells of unequal sizes, known as the stalked and swarmer cells (Figure 1A).

In this study, we performed high-precision single-cell time-lapse microscopy studies on *C. crescentus* and *E. coli* to uncover the intrinsic principles of bacterial cell size homeostasis for both symmetric and asymmetric divisions.

RESULTS AND DISCUSSION

C. crescentus Does Not Appear to Vary Its Cell Length in Response to Changes in Nutrient Availability, Cell Density, or Temperature

In this study, we focused on cell length, as this is the cell size dimension that changes during the cell cycle. Cell length in bacteria is generally thought to be sensitive to environmental conditions. For example, *E. coli* and other bacteria are known to modulate their average length in response to nutrient availability and cell density (Akerlund et al., 1995; Schaechter et al., 1958). Using high-precision image analysis software (Sliusarenko et al., 2011), we found that, although the cell width slightly varies (Harris et al., 2014), the length of asymmetrically dividing *C. crescentus* appears insensitive to environmental fluctuations. There was no discernable difference in cell length distributions between *C. crescentus* populations grown in nutrient-poor (M2G) or nutrient-rich (PYE) medium (Figure 1B) despite significant differences in growth rates, with doubling times of 146 ± 5 min in M2G (mean \pm SD, $n = 2$ experiments) and 96 ± 1 min in PYE ($n = 2$). Varying cell density ($OD_{660nm} < 0.3$ versus > 0.8) or temperature (25°C versus 30°C) also had little effect on cell length distributions (Figure 1B). These distributions were repro-

ducible from day to day (data not shown). Thus, *C. crescentus* populations exhibited the same cell length distributions under all growth conditions tested, indicating robust cell length homeostasis. This allowed us to examine the intrinsic properties of cell length control, without concern of interference from environmental fluctuations.

C. crescentus Controls Its Cell Length

Studying cell length control requires precise measurements of cell length over at least one full cell cycle. Obtaining swarmer cells in G1 phase is relatively easy (Evinger and Agabian, 1979). However, the synchronization technique does not distinguish between the “young” swarmer cells that are fresh from division and the “older” swarmer cells that are about to become stalked cells. Furthermore, the technique, like most cell-cycle synchronization methods, perturbs cellular metabolism. Therefore, we performed time-lapse phase contrast microscopy of asynchronous populations and monitored growth and division (see Experimental Procedures). Analysis showed that, when cells were spotted on regular 1% agarose pads containing M2G medium, the average cell length became significantly shorter ($\sim 10\%$) even after a single cell cycle. We reasoned that this cell shortening was likely caused by physical constraints due to immobilization on the solid agarose (1%) substrate. To reduce this potential “sticky” problem, we spotted cells on soft agarose (0.3%) pads. In this more aqueous environment, cells were more loosely immobilized, allowing newborn swarmer cells to swim away immediately after their physical separation from the stalked cell sibling following division (Movie S1 available online). This soft-agarose microscopy set-up allowed us to track stalked cells over time and to measure their length from birth (L_b) to division (L_d). In this environment, the cell lengths remained similar between divisions, with cell length at birth $L_b = 2.43 \pm 0.39 \mu\text{m}$ (mean \pm SD, $n = 252$ cells) after the first division versus $L_b = 2.42 \pm 0.40 \mu\text{m}$ ($n = 193$ cells) after the second division. Because swarmer cells occasionally became immobilized on the soft-agarose surface after a short swim (Movie S1), we also obtained measurements for swarmer cells, although many fewer (see Experimental Procedures).

As expected (Terrana and Newton, 1975), division was asymmetric, with a division ratio DR (length between the stalked pole and the division site divided by the total length) of 0.56 ± 0.04 ($n = 706$ cells) (Figure 1C). The populations of newborn stalked and swarmer cells were characterized by narrow distributions of cell length at birth (Figure 1D). Both cell types had a very similar relative variability in length, as measured by the coefficients of variation (CV, standard deviation/mean) of 16% and 18% for stalked and swarmer cells, respectively. Despite the shift in average length at birth, swarmer and stalked cells displayed similar distributions of cell lengths at division, indicating that swarmer cells must grow more than stalked cells to compensate for their shorter size at birth. The cell length distributions at division were narrow, with a CV of about 12% (Figure 1D). This value is comparable to the CV determined for *Saccharomyces cerevisiae* at budding (CV = 17%) and *Schizosaccharomyces pombe* at fission (CV = 6%) (Di Talia et al., 2007; Lord and Wheals, 1981; Svecizer et al., 1996), both of which are known to display cell size regulation.

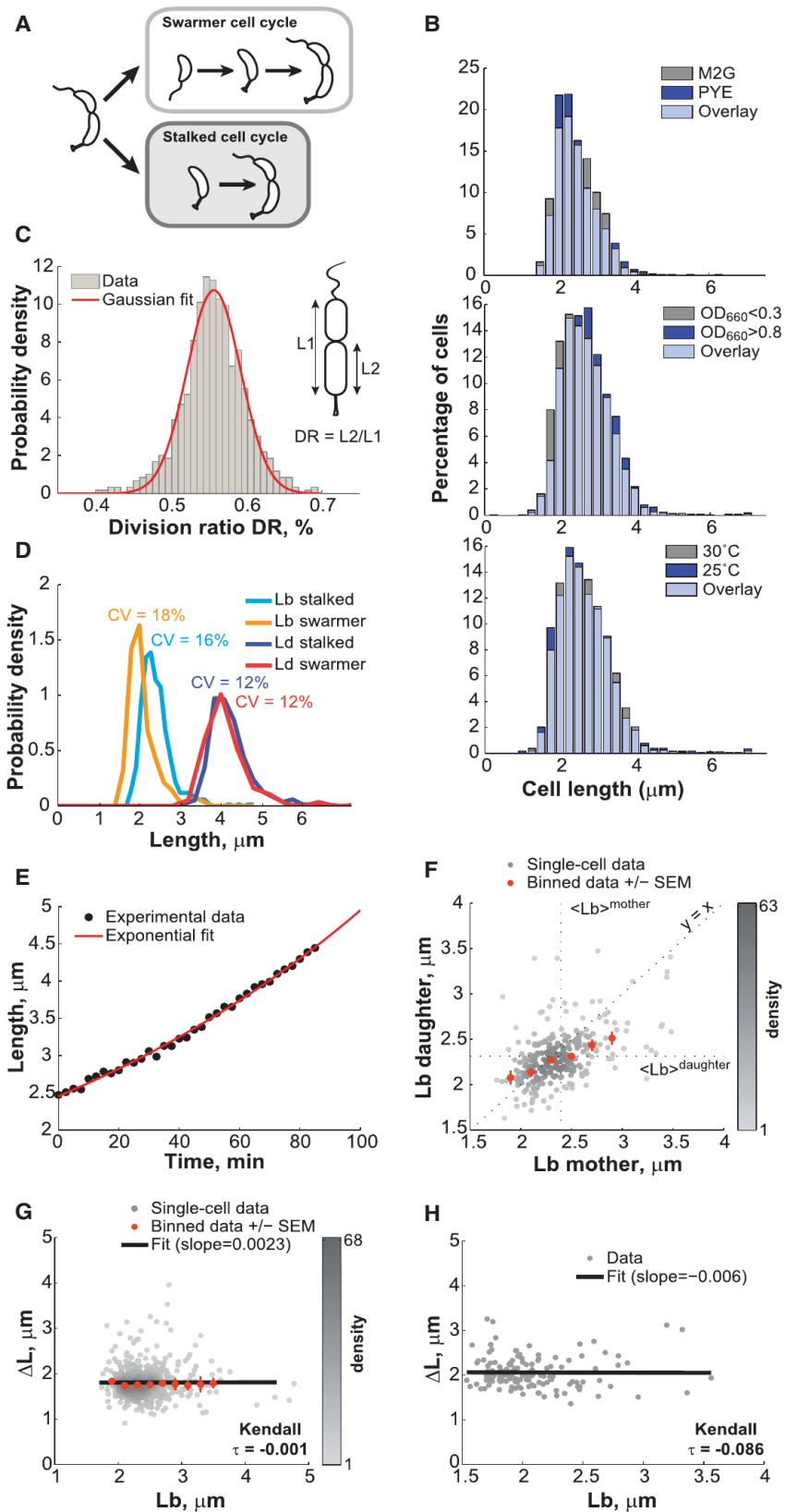


Figure 1. Cell Length Control in *C. crescentus*

(A) Schematic of the dimorphic cell cycle of *C. crescentus*. Each division generates two different progeny: the smaller swarmer cell and the longer stalked cell. The swarmer cell is motile thanks to a polar flagellum until the cell transitions to a stalked cell.

(B) Distribution of *C. crescentus* CB15N lengths under different environmental conditions. Cells were grown in M2G at 30°C and were imaged on 1% agarose pads at an $OD_{660nm} < 0.3$ except if stated otherwise.

(C) Histogram representing the distribution of the division ratio DR for *C. crescentus* CB15N cells ($n = 706$) grown in M2G medium and imaged on 0.3% agarose pads.

(D) Distribution of cell length at birth Lb and at division Ld for stalked cells ($n = 565$) and swarmer cells ($n = 141$) grown in M2G medium at 30°C and imaged at $OD_{660nm} < 0.3$ on 0.3% agarose pads. The coefficient of variation (SD/mean) for each distribution is shown.

(E) Representative growth curve of a single stalked cell (black circles) grown in M2G medium at 30°C on a 0.3% agar pad. The red line is the best fit of the data with an exponential function.

(F) Partial inheritance of Lb from one generation to the next ($n = 457$ stalked cells). The vertical and horizontal dotted lines indicate the mean length at birth for mother and daughter cells. The line $y = x$ is also plotted for comparison purposes.

(G) Dependence of the elongation over a cell cycle (ΔL) on Lb for stalked cells. Gray dots represent single-cell data, whereas orange dots represent the average of binned data \pm SEM. The shade of gray represents the density of points in a given area of the graph. The black line represents the linear fit to the single-cell data.

(H) Dependence of ΔL on Lb for swarmer cells. There were not enough cells ($n = 141$) to bin the data. See also Figures S5 and S7 and Movie S1.

Cell Size Compensation Is Partial over a Single Generation

Further analysis was primarily done on the stalked progeny, given their higher sample size. Their cell elongation was consistent with exponential growth (Figure 1E), as reported previously (Siegal-Gaskins and Crosson, 2008). Exponential growth implies that a cell size compensation mechanism must be at work to maintain the narrow cell length distributions that we observed. We indeed found that stalked cells born shorter than the population average produced stalked cells that were comparatively longer than their mothers (L_b of daughter > L_b of mother) (Figure 1F). The reverse was true for cells born longer than the mean; their progeny were comparatively shorter. However, the compensation was only partial (Figure 1F). This was surprising because a sizer-based model (with or without timer) implies that all cells shorter than the critical size grow until they reach their size threshold. As a result, there should not be any correlation in cell length between mothers and daughters for cells born shorter than the critical size. In other words, the “short” phenotype is not an inheritable feature when a critical size mechanism is in place, unlike what we observed.

C. crescentus Cells Elongate by a Constant Amount on Average, Irrespective of Cell Length at Birth

Another key characteristic of any sizer-based model is that cells born smaller than the mean size grow, on average, more before dividing than cells born longer. Thereby, in a sizer model, cell extension during the cell cycle ($\Delta L = L_d - L_b$) displays a strong negative correlation with the cell length at birth (L_b) for cells born shorter than the critical size (Fantès, 1977; Sveiczer et al., 1996). Strikingly, we found no significant correlation (slope ~ 0 , Kendall $\tau = -0.001$) between ΔL and L_b for stalked cells ($n = 565$, Figure 1G). This seemed to be also true for the swarmer progeny despite lower statistics ($n = 141$, Figure 1H). These data suggest that *C. crescentus* cells do not sense a certain size to regulate their length. Instead, they simply elongate the same amount on average ($\Delta L = 1.81 \pm 0.36 \mu\text{m}$ and $2.06 \pm 0.35 \mu\text{m}$ for stalked and swarmer progeny, respectively) before dividing, regardless of their size at birth.

E. coli Does Not Sense a Cell Size Threshold to Control Its Length

The surprising lack of cell size threshold in *C. crescentus* prompted us to revisit the critical size paradigm in *E. coli*, which had mostly been inferred from population studies under fast-growing conditions (Cooper, 1991; Cooper and Helmstetter, 1968; Donachie, 1968). Because the size of *E. coli* is sensitive to changes in nutrient availability or cell density, it was crucial to maintain constant growth conditions during measurements. For this, we used a microfluidic device (Ullman et al., 2013) that allowed us to track hundreds of *E. coli* BW25113 cells at the high temporal resolution of 5 s for hours (Figure S1 and Movie S2). We used fast-growth conditions (LB-rich medium at 30°C) that resulted in an interdivision time of 27 ± 5 min (mean \pm SD, $n = 1,305$ cells). We verified that the growth rate and the average cell length at birth remained constant through the entire 7 hr experiment (Figures 2A and 2B), indicating steady-state conditions. We also verified that the position of the cells in the microfluidic chamber had no influence on

these parameters (Figure S1B). Cell elongation at the single-cell level was well approximated with an exponential function (Figure 2C), consistent with exponential growth.

Cell lengths at birth and division were narrowly distributed, with low CV of 12% and 11% (Figure 2D), respectively, consistent with previous reports (Koppes et al., 1980; Wakamoto et al., 2005). Strikingly, *E. coli*, even under fast-growing conditions, behaved similarly to *C. crescentus* in many respects. First, cells born shorter or longer than the mean only displayed partial cell size compensation over a single cell cycle (Figure 2E). Second, there was no correlation between the amount of elongation over a cell cycle and the length of cells at birth (Figure 2F). *E. coli* cells grew, on average, the same length ($\Delta L = 3.23 \pm 0.60 \mu\text{m}$, mean \pm SD, $n = 1,305$ cells) before dividing, independent of their initial size. These results demonstrate a fundamental conflict with the predictions of a sizer-based mechanism.

By analyzing published microfluidic data (Wang et al., 2010) (see Experimental Procedures), we found that ΔL is also independent of L_b for *E. coli* strains MG1655 and B/r growing at 37°C in 1 μm wide linear chambers (Figure S2), suggesting that our observations are independent from the strain, the microfluidic chamber geometry, and the temperature.

A Constant Cell Elongation per Cell Cycle Results in Cell Size Homeostasis

How can cells control their size in the absence of a sizer mechanism? Our data show that, although there are significant fluctuations in ΔL values, cells elongate, on average, a constant amount before dividing, irrespective of their length at birth (Figures 1G, 1H, 2F, and S2). A constant elongation—or the addition of a volume increment—can, at least theoretically, lead to cell size homeostasis (Amir, 2014; Voorn and Koppes, 1998). Figure 3A shows schematically how a constant length extension followed by a symmetric division can compensate for cell size fluctuations within a few generations. Mathematically (see Extended Experimental Procedures, Equation 3), cell lengths at birth that deviate from the mean converge exponentially over generations toward the constant elongation value in the absence of noise (Figure 3B). This assumes that all cells extend by the exact same increment ΔL , divide precisely in half, and grow at an identical relative rate α . In reality, these parameters fluctuate around an average value. When the experimental variability in ΔL , DR , and α was considered in our mathematical model, stochastic simulations (see Extended Experimental Procedures) showed that virtual newborn cells of varying initial lengths (1–10 μm) converge to the expected steady-state distribution of L_b within two to four divisions (Figure 3C). After that, L_b fluctuates around the average value (Figure 3D) and the population as a whole reaches a steady-state regime in which the L_b distribution remains stable over time (Figure 3E). Thus, the constant extension mechanism does not need to be precise to be robust.

Remarkably, the steady-state distribution of L_b obtained from the simulations almost perfectly overlapped with the experimental distribution (Figure 3F), providing further support for the constant extension model. The model also quantitatively predicts the cell size compensation profile after one generation (Figure 3G compared to 2E), as well as over subsequent generations (Figure 3H compared to 3I). Because we were able to track cell

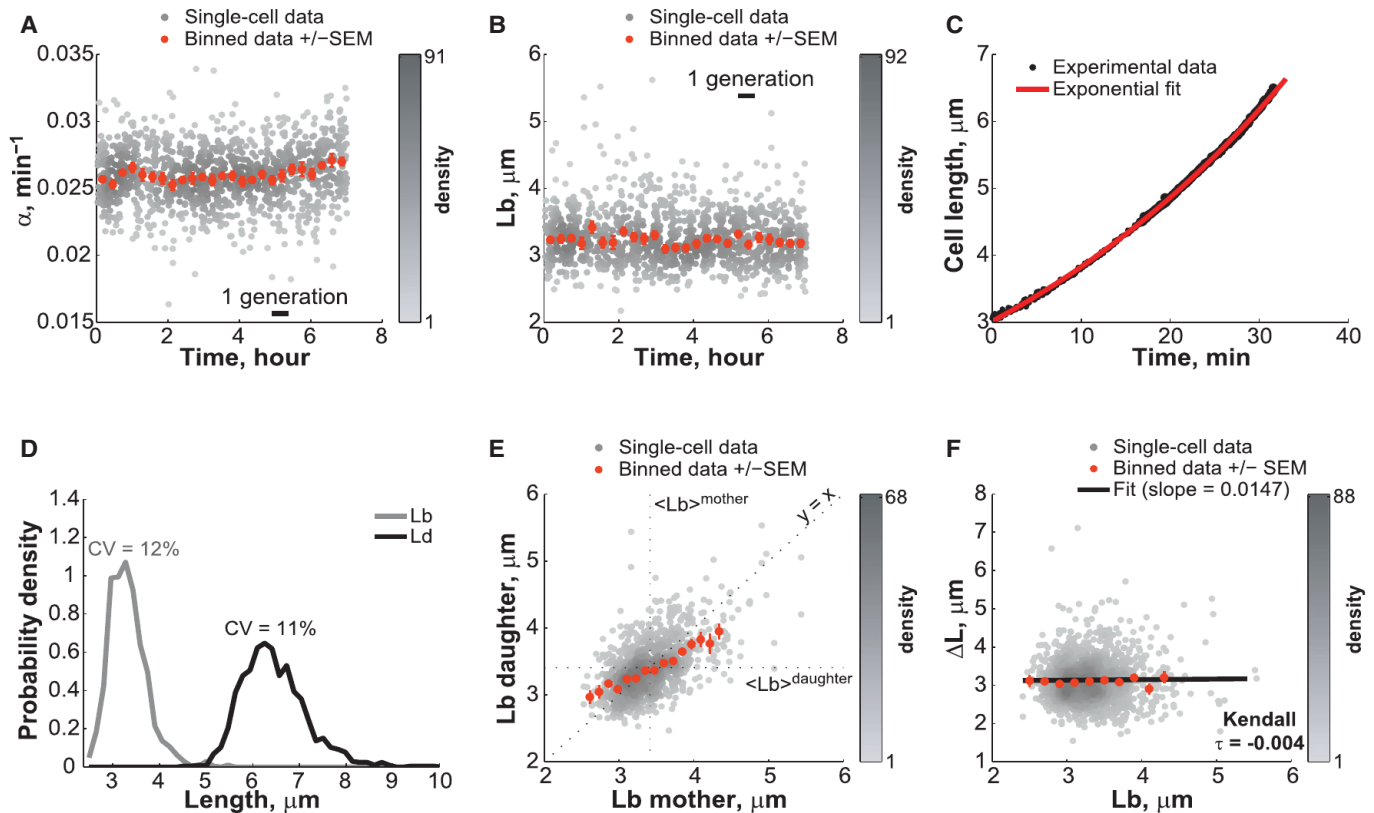


Figure 2. Steady-State Growth of *E. coli* BW25113 Cells in Microfluidic Chambers

E. coli BW25113 cells ($n = 1,305$) were grown in microfluidic chambers at 30°C in nutrient-rich LB medium.

(A) Relative growth rate α over the course of the > 7 hr long microfluidic experiment. Gray dots represent single-cell data, whereas orange dots represent the average of binned data \pm SEM.

(B) Same as (A) except that L_b was plotted instead of α .

(C) Representative growth curve of a single cell (black circles). Length was measured every 5 s. The red line is the best fit of the data with an exponential function.

(D) Distribution of L_b and L_d .

(E) Partial inheritance of L_b from one generation to the next.

(F) Dependence of ΔL on L_b . The black line represents the linear fit to the single-cell data.

See also Figures S1, S2, S5, S7 and Movie S2.

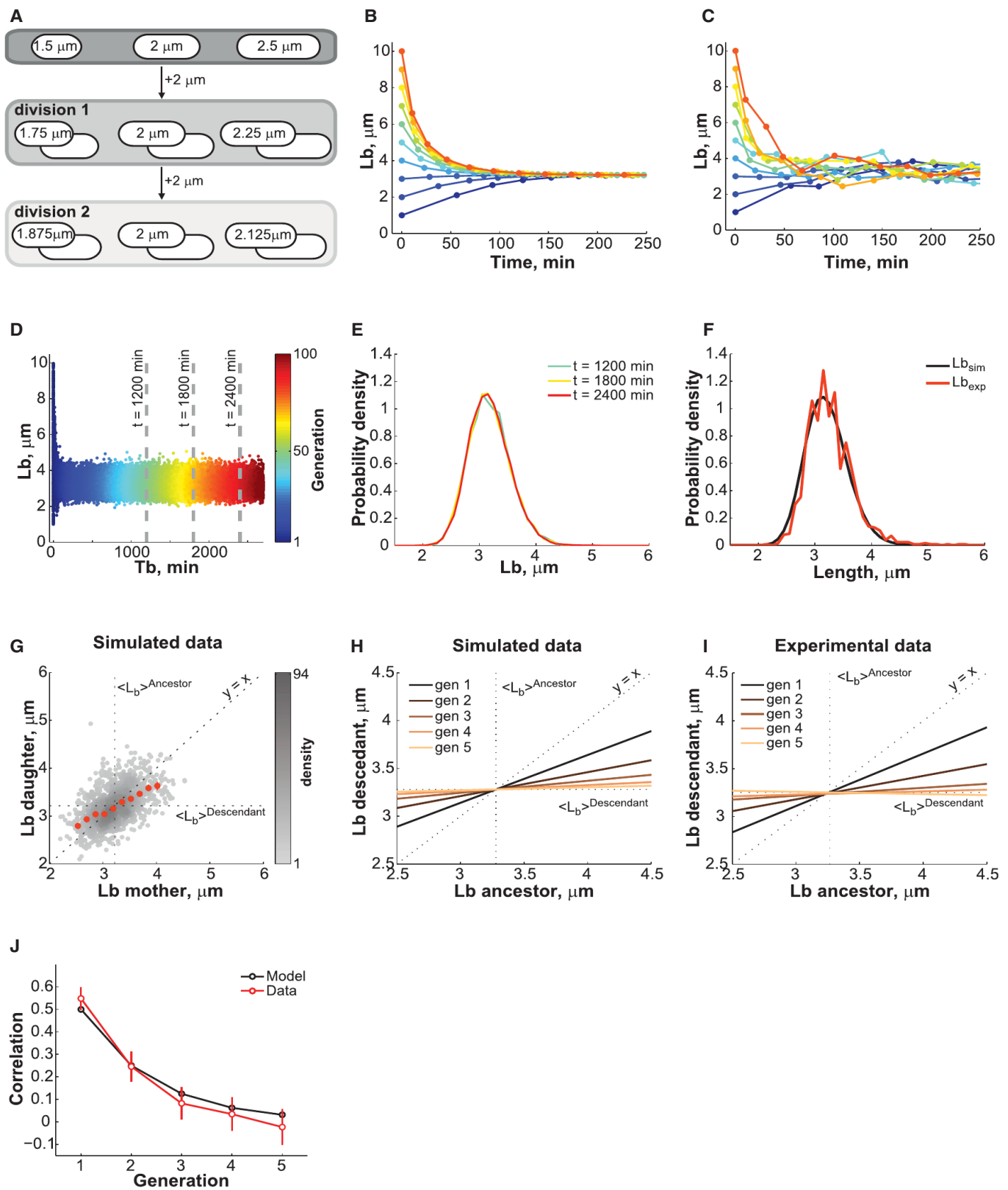
lineages over several generations in our *E. coli* microfluidic experiments, we also showed that the correlation in L_b between ancestors and descendants drops with the number of generations (Figure 3J), following an exponential decay consistent with the constant extension model (see Extended Experimental Procedures, Equation 3). Thus, the model accurately predicts the degree of correlation in L_b between ancestors and descendants.

The constant extension model works equally well for asymmetric divisions. Using the experimental ΔL and DR values for the *C. crescentus* data set, simulations accurately reproduced the experimental L_b distribution of both stalked and swarmer daughter cells (Figure S3A). For symmetrically dividing bacteria (average division ratio $\langle DR \rangle = 0.5$), the average length at birth ($\langle L_b \rangle$) equals the average elongation ($\langle \Delta L \rangle$), as shown for *E. coli* (Figure S3B). For asymmetrically dividing bacteria, this is not the case (Figure S3C). $\langle \Delta L \rangle$ of the longer progeny will correspond to $\langle L_b \rangle$ of the smaller one and vice versa (see Extended Experimental Procedures, Equation 2). For example,

the average length extension of *C. crescentus* stalked cells is, within the measurement error, equal to the length of swarmer cells at birth (Figure S3C).

Precision in Cell Length at Birth Is Dictated by the Division Ratio and the Length Extension between Divisions

Just as the average L_b value is determined by the average ΔL and DR values, deviations from this targeted size—which is reflected by the CV of the L_b distribution—will be dictated by the precision of both the constant extension mechanism and the positioning of the division machinery. Note that, although the average ΔL is equal to the average L_b for symmetrically dividing cells, variations in ΔL and L_b can be different (Figure S3B). The constant extension model (see Extended Experimental Procedures, Equations 14 and 15) quantitatively predicts how the CV of L_b depends on the CV of ΔL and DR (Figures S3D and S3E). We found that the predicted CV values of L_b are in excellent agreement with the experimental values (Figures S3D and S3E).



(legend on next page)

Cell-Cycle Time Increases with Shorter Cell Length at Birth

How do cells elongate by the same amount on average? Is it by modulating their cell-cycle time T or by changing their relative elongation rate α ? We found that, for both the *E. coli* and *C. crescentus* data sets, T decreases as L_b increases (Figures 4A and 4B), whereas α remains fairly constant (Figures 4C and 4D), as previously reported for *E. coli* (Osella et al., 2014). Note that the *relative* elongation rate α (min^{-1}) corresponds to the *relative* length increase over time, not to be confused with the *absolute* elongation rate ($\mu\text{m min}^{-1}$), which is the absolute increase of cell length over a period of time. By virtue of their exponential growth, a constant relative elongation rate implies that the absolute elongation rate averaged over the cell cycle will increase with increasing L_b , which is what we observed (Figures 4C and 4D).

Collectively, our data show that cells modulate their cell-cycle time, and not their relative growth rate, to achieve the same length extension. As a result, the so-called “normalized cell-cycle time” αT is negatively dependent on L_b (Figures 4E and 4F). This negative dependence is often taken as supportive evidence for a sizer-based mechanism. The rationale for this is that, if a cell needs to reach a certain size before committing to division, shorter cells at birth require longer cell-cycle times. However, this negative dependence between αT and L_b is also expected from the constant extension model. In fact, if we used the average ΔL values obtained from the *E. coli* and *C. crescentus* stalked cell experiments ($\Delta L = 3.23 \mu\text{m}$ and $\Delta L = 1.81 \mu\text{m}$, respectively), we found that the analytical expression derived from the constant extension model (see Extended Experimental Procedures, Equation 5) describes the averaged data very well (Figures 4E and 4F).

The Constant Length Extension between Divisions Changes with Nutrient Availability in *E. coli*

It is well known that *E. coli* changes its size in response to nutrient availability (Schaechter et al., 1958). If cell size homeostasis works through a constant elongation irrespective of the composition of the growth medium, we would expect the fixed cell length extension in nutrient-poor medium to be smaller than in the nutrient-rich LB medium, with the average ΔL value

matching the average L_b value for each growth medium. To test this hypothesis, we grew *E. coli* BW25113 in microfluidic chambers with M9-supplemented medium (M9 salts supplemented with 0.1% casamino acids and 0.2% glucose) instead of LB medium. Under steady-growth conditions (Figure S4), the cell-cycle time in M9 supplemented medium was longer than in LB medium with $T = 42 \pm 12 \text{ min}$, and the cell lengths at birth and at division were shifted to lower values (Figure 5A), with $L_b = 2.32 \pm 0.38 \mu\text{m}$ and $L_d = 4.59 \pm 0.71 \mu\text{m}$. Cell size control in M9 medium remains precise, with CV values of 17% and 15% for L_b and L_d , respectively.

Importantly, as in rich medium, the correlation for the length at birth between mothers and daughters was 0.5 (Figure 5B), as expected from the constant extension model. In addition, ΔL and L_b remained uncorrelated (Kendall $\tau = 0.03$, Figure 5C). Curiously, we observed correlations between some parameters (α of daughter versus α of mother and ΔL versus α , Figures S5A and S5B) in the M9 data that were absent in the LB data set (Figures S5C and S5D); however, these correlations had no impact on cell size regulation (Figures 5D, S5E, and Extended Experimental Procedures). Consistent with our hypothesis, cells elongated by a shorter increment before division occurs, with $\Delta L = 2.26 \pm 0.59 \mu\text{m}$ in M9-supplemented medium compared to $3.23 \pm 0.6 \mu\text{m}$ in rich LB medium to set the corresponding length at birth. This suggests that cells modulate their average length by changing the average ΔL in response to changes in nutrient availability.

If the latter is true, we reasoned that mutants impaired in the transport of the carbon source from the environment may behave as if they were growing in a carbon-poor medium: cells would implement a smaller average ΔL while fully maintaining size homeostasis (i.e., similar $\text{CV}_{\Delta L}$ as wild-type). The phosphotransferase system (PTS) is involved in the efficient transport of sugars across the membrane. Deletion of genes encoding the core PTS components PtsH and PtsI in *E. coli* resulted in a reduced growth rate (>2-fold) and a visibly shorter cell morphology in glucose-containing M9 supplemented medium (Figures 5E and 5F). The ΔptsH and ΔptsI mutants were also slightly thinner (data not shown). Consistent with our expectation, quantitative analysis (see Extended Experimental Procedures) showed that

Figure 3. Cell Size Control by a Constant Extension Model

- (A) Schematic showing how a constant elongation allows cells of below- and above-average length to recover over time and generations.
- (B) Analytical dependence of L_b on time for ten cells with L_b ranging from 1 to 10 μm . The time between data points represents the cell cycle (generation) time for that particular cell. See Table S1 for input parameter values.
- (C) Same as (B) except that the experimentally observed fluctuations around the DR , α , and ΔL mean values were added to the simulation.
- (D) Evolution of L_b for 500 simulated cells with initially wide distribution of lengths.
- (E) Plot showing the distribution of L_b at the three time points indicated by the gray dashed lines in (D). All cells present at the indicated times were considered to build the L_b distributions.
- (F) Plots showing the L_b distributions obtained from experiment and simulation (see Table S1 for input parameter values).
- (G) Plot showing the dependence in L_b between mother and daughter cells. The data were obtained from a stochastic simulation of the constant extension model using the same number of cells ($n = 1,305$) as in the experiment shown in Figure 2. Gray dots represent single-cell data, whereas orange dots represent the average of binned data $\pm \text{SEM}$.
- (H) Degree of inheritance of L_b over one to five generations in simulations. Each line represents the best linear fit to the single-cell data of L_b for an ancestor and its descendant from the first to fifth generation.
- (I) Same plot as in (H) except that the data were derived from the *E. coli* LB experiment.
- (J) Evolution of the correlation between L_b from ancestors to descendants over generation expected from the model (black) or observed experimentally (red). The correlation was based on linear regressions of single-cell data. The error bars represent the 95% confidence interval around the slope value.
- See also Figures S3 and S5.

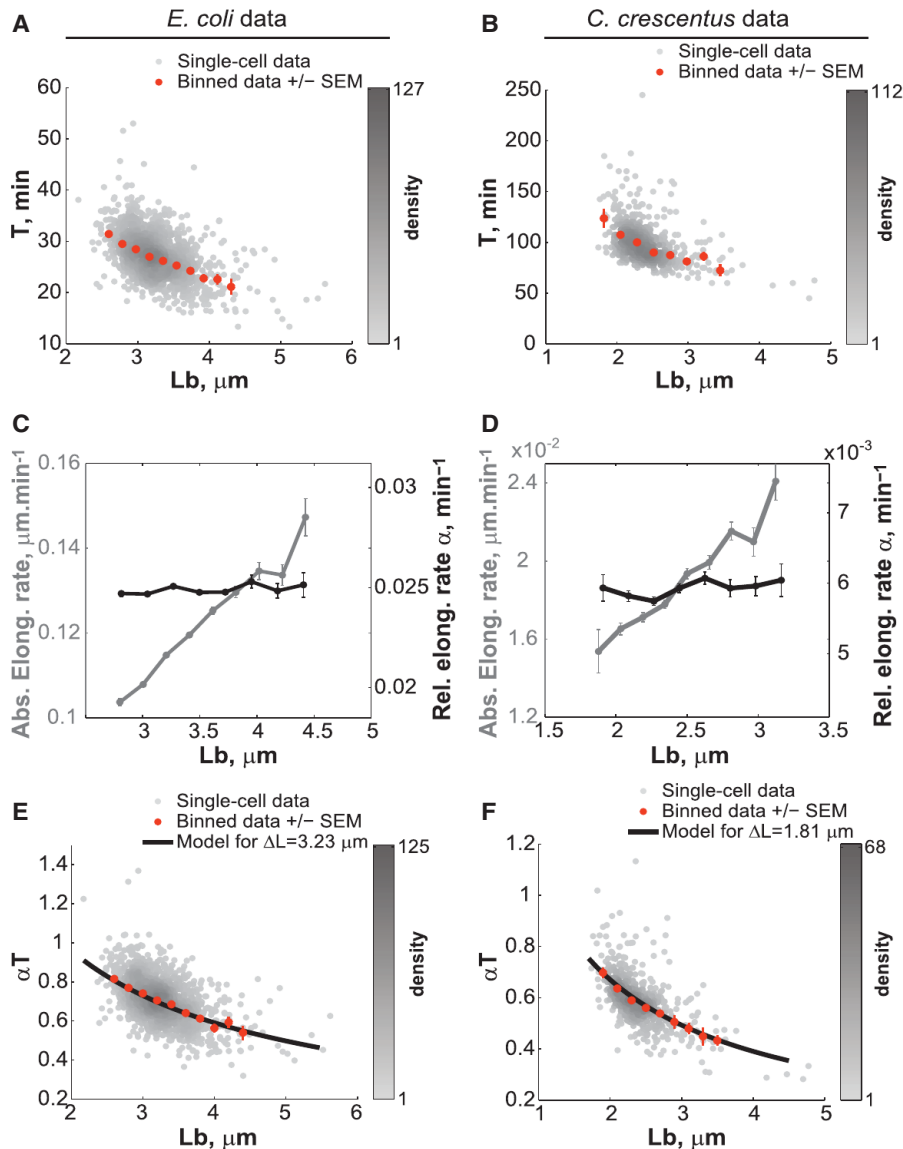


Figure 4. Cells Correct Their Length at Birth by Modulating Their Cell-Cycle Time and Not Their Relative Growth Rate

(A) Dependence of the cell-cycle time T on L_b for *E. coli* grown in LB medium ($n = 1,305$ cells). (B) Same as (A) but for *C. crescentus* stalked cells grown in M2G medium ($n = 565$ cells). (C) Dependence of the absolute (gray) and relative (black) elongation rate α on L_b for *E. coli*. (D) Same as (C) but for *C. crescentus*. (E) Dependence of the normalized cell-cycle time αT on L_b for *E. coli*. The black line represents the expected average αT from the constant extension model, as described by the equation $\alpha T = \ln(1 + \Delta L / L_b)$. The black line was drawn using the average ΔL value measured experimentally. (F) Same as (E) but for *C. crescentus*. Error bars represent \pm SEM.

and S5G, with $\tau = 0.10$ and 0.03 for *E. coli* in M9 and LB media and with $\tau = 0.06$ for *C. crescentus*, respectively). In other words, the accuracy of a cell in implementing the targeted ΔL has no bearing on the precision of its descendants.

Could the Constant Length Extension Be Applied at a Cell-Cycle Event Other Than Division?

So far, we have considered the case of a constant elongation being applied from one division to the next (Figure 6A). But what if the point of control for cell size homeostasis occurs at an earlier cell-cycle event X, and completion of this event X triggers division after a constant amount of time, or timer δt , has elapsed? In this scenario, the constant elongation (referred to as ΔL^* to distinguish it from the elongation ΔL that we actually

measured between consecutive divisions) would be applied at this event X, and not at division. For example, this early control event could be the initiation of DNA replication, as theoretically proposed before (Amir, 2014). This is an important consideration, as DNA replication initiation is often assumed to control division. In a scenario in which ΔL^* is applied at cell-cycle event X, the division cycle would be shifted out of phase relative to the constant elongation cycle, and division would follow the completion of a fixed elongation ΔL^* after a timer δt (Figures 6B and 6C). Could this scenario account for the experimental data we obtained?

The Constant Extension Mechanism Has No Memory

The constant extension mechanism is precise, but not perfect, as reflected by the CV of ΔL (Figures S3D and S3E). As such, ΔL values fluctuate around an average value (i.e., the targeted value) from cell to cell and from generation to generation. Importantly, the microfluidic data showed that the constant extension mechanism has no memory, as the ΔL values of mother and daughter cells were not significantly correlated (Figures S5F

and S5G, with $\tau = 0.10$ and 0.03 for *E. coli* in M9 and LB media and with $\tau = 0.06$ for *C. crescentus*, respectively). In other words, the accuracy of a cell in implementing the targeted ΔL has no bearing on the precision of its descendants.

First, we considered the case in which the timer δt is shorter than the interdivision time T (Figure 6B). This would illustrate the *C. crescentus* situation ($T = 100$ min) in which DNA replication initiates early in the cell cycle and completes before division (McAdams and Shapiro, 2009). Simulations of the phase-shifted model showed significant discrepancies with the experimental

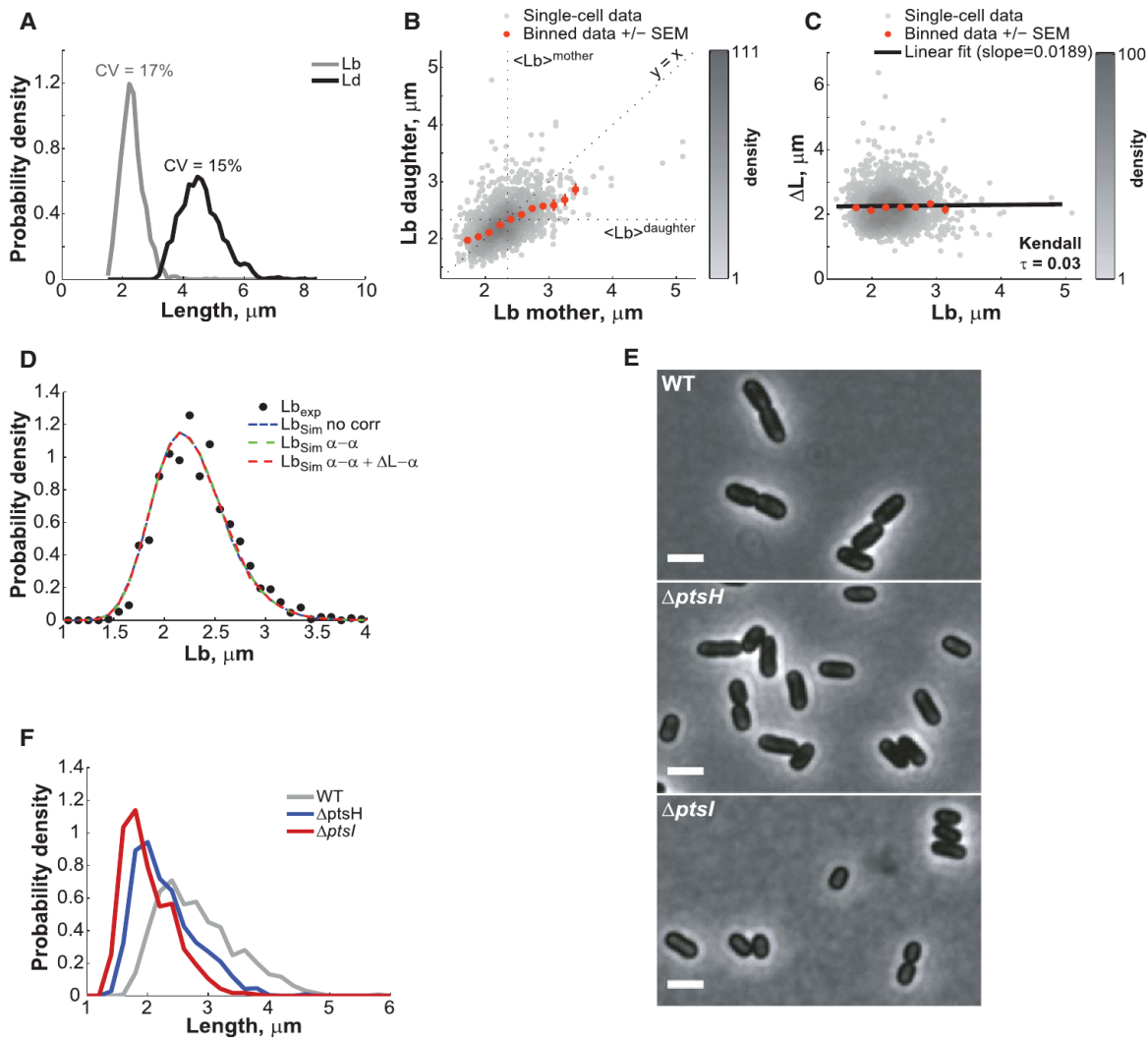


Figure 5. *E. coli* Also Controls Its Length through a Constant Elongation in M9-Supplemented Medium

(A) Distribution of L_b and L_d of BW25113 cells ($n = 1,528$) grown in M9-supplemented medium at 32°C in a microfluidic chamber over 6 hr.

(B) Partial inheritance of L_b from one generation to the next.

(C) Dependence of ΔL on L_b . The black line represents the linear fit to the single-cell data.

(D) Experimental and simulated L_b distributions. “No corr,” “ α - α ,” and “ α - α + ΔL - α ” refer to the type of correlations included in the simulations (as described in the Extended Experimental Procedures).

(E) Phase-contrast images of BW25113 wild-type, $\Delta pstH$, and $\Delta ptsI$ grown in M9-supplemented medium in liquid cultures and spotted on 1% agarose pads. Scale bars, 2 μm .

(F) Cell length distributions of wild-type, $\Delta pstH$, and $\Delta ptsI$ populations.

See also Figures S4 and S5.

C. crescentus data for any timer that starts, on average, earlier than at 80% of the division cycle (i.e., for any timer $\delta t > 20$ min for $T = 100$ min). For example, in the phase-shifted model, ΔL (cell length extension between consecutive divisions) and L_b show a significant negative correlation (Figure 6D), in contrast to what was observed experimentally (Figure 1G). Other relationships between variables were also inconsistent with the experimental data (Figures S6A–S6C). Therefore, a constant elongation is unlikely to be applied at DNA replication initiation—or at any other early cell-cycle event—to control cell size homeostasis in *C. crescentus*. Whether division was asymmetric (Figures 6D

and S6A–S6C) or symmetric (Figures S6D–S6F) did not alter the conclusion.

We also considered the relevant case in which the timer δt would be longer than the interdivision time T (Figure 6C). Previous work has proposed that, in *E. coli*, DNA replication initiation and cell division are separated by a constant timer of about 60 min (Cooper and Helmstetter, 1968), which would exceed the doubling times of 27 and 42 min that we observed for *E. coli* growing in LB and M9 growth media, respectively. To consider these fast-growth cases, we ran simulations of the phase-shifted model with timers longer than the interdivision

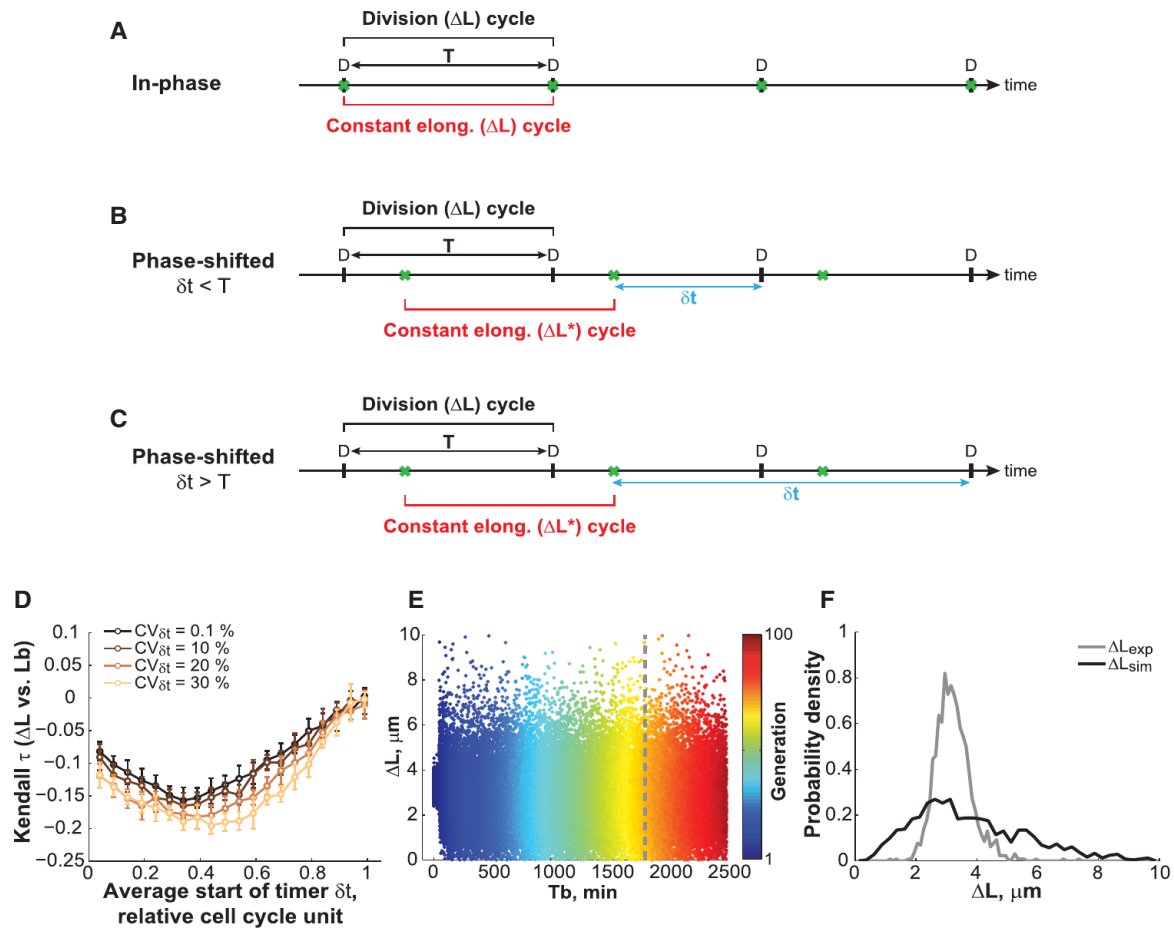


Figure 6. Testing the Phase-Shifted Constant Extension Model

(A) Schematic of the constant extension model in which the division cycle (black bracket) is in phase with the constant elongation cycle (red bracket). The event under cell-cycle control (green crosses) is division D.

(B) Schematic of the “phase-shifted” model with a timer $\delta t < T$. In this model, the elongation increment (ΔL) is applied to a cell-cycle event X (green crosses) that is coupled to division by a timer δt shorter than the interdivision time T .

(C) Schematic representation of the “phase-shifted” constant extension model with $\delta t > T$ and with the constraint of a single event X per division cycle (see Extended Experimental Procedures and Table S1).

(D) Simulations of the “phase-shifted” model with a timer $\delta t < T$ (see Extended Experimental Procedures and Table S1). The Kendall correlation coefficient between L_b and ΔL depends on how early in the cell cycle the timer δt starts. Shown is the mean \pm SD of ten simulations performed with 1,500 cells for $DR = 0.56$.

(E) Simulations of the “phase-shifted” model with a timer $\delta t > T$. Shown is a scatter plot of ΔL (from birth to division) for 500 simulated cells at each generation.

(F) Same as (E) but plot showing the distribution of ΔL at the time indicated by the dashed line in (E).

See also Figure S6.

times. They resulted in the generation of widely abnormal cell size distributions (Figures S6G and S6H), which arose from the fluctuation in number of event X (e.g., DNA replication initiation) occurring between two divisions. Even when event X was forced to happen exactly once per division cycle (by adding constraining rules to the model, see Extended Experimental Procedures) to reduce the L_b variability, the ΔL values remained aberrantly variable (Figures 6E and 6F). Furthermore, the cell length extension between divisions (ΔL) was correlated between mothers and daughters in the simulated data (Figure S6I); that is, this phase-shifted model with $\delta t > T$ displays memory because the constant elongation ΔL^* overlaps with two consecutive division cycles (Figure 6C). This correlation in ΔL between mothers and

daughters is in contradiction with the experimental data (compare Figures S5F and S5G with Figure S6I).

Collectively, these experimental results are inconsistent with the hypothesis that the constant elongation is applied at the initiation of DNA replication or any cell-cycle event that occurs within the first 80% of the interdivision time. This also excludes the formation of the FtsZ cytokinetic ring, implying that this event is unlikely to dictate the timing of cell division. This is in agreement with single-cell observations that FtsZ ring formation and cell constriction are uncorrelated in time (Tsukanov et al., 2011). Our analysis suggests that the cycles of constant elongation and division are in phase or are close to it (Figure 6A), indicating that a late cell-cycle stage is the control point. For instance, a late

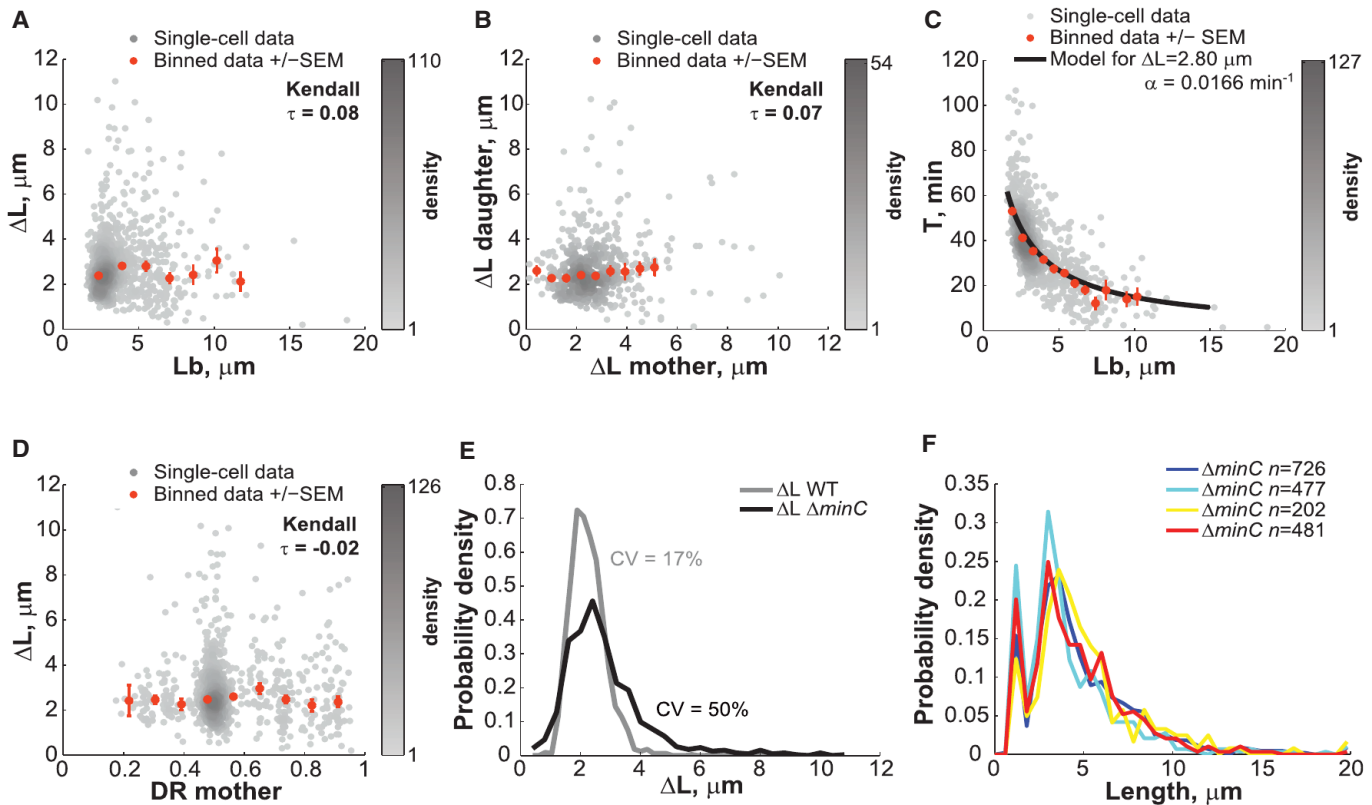


Figure 7. The Constant Extension Mechanism Is Robust with Respect to the Cell Length and Division Site Placement

E. coli BW25113 $\Delta minC$ cells were grown in microfluidic chambers at 30°C in M9-supplemented medium over 5.5 hr.

(A) Dependence of ΔL on L_b ($n = 959$ cells). Gray dots represent single-cell data, whereas orange dots represent the average of binned data \pm SEM.

(B) Degree of inheritance of ΔL over one generation ($n = 510$ cells).

(C) Dependence of the cell-cycle time T on L_b ($n = 959$ cells). The black line represents the expected average αT from the constant elongation model, as described by the equation $T = \ln(1 + \Delta L/L_b)/\alpha$. The black line was drawn using the average ΔL and α values measured experimentally.

(D) Dependence of ΔL on DR .

(E) Distributions of length extensions for wild-type and $\Delta minC$ cells.

(F) Cell length distributions of BW25113 $\Delta minC$ cells grown in separate 2 ml liquid cultures in M9-supplemented medium.

See also Movie S3.

cell-cycle event (e.g., a late step of chromosome segregation potentially sensed by FtsK) may communicate with the cell division machinery to trigger FtsZ ring constriction. Division would then restart a cycle of constant length extension.

The Constant Extension Mechanism Is Robust with Respect to Cell Length and Division Positioning

In the constant extension model, cell size homeostasis is based on a simple governing principle: cells trigger division once they have elongated the targeted ΔL (\pm noise). Does the simple rule of adding a constant length still apply when cells are aberrantly long or when division is misplaced? To address this question, we carried out microfluidic experiments with the *E. coli* $\Delta minC$ mutant (Movie S3). Without MinC, the min system that regulates the precision of FtsZ ring placement is defective; as a result, cells divide not only at midcell, but also at polar, DNA-free regions, resulting in the appearance of minicells (Adler et al., 1967). As expected, the large imprecision in division placement in the $\Delta minC$ mutant leads to very wide distribution of cell lengths at birth (CV

of $L_b = 52\%$). But despite these aberrations, $\Delta minC$ cells elongated a constant amount between divisions, regardless of their length at birth (Figure 7A, Kendall $\tau = 0.08$; anucleate minicells were excluded from the analysis, as they do not grow). Even very long cells grew, on average, the same amount as short cells before dividing (Figure 7A). Thus, the constant extension mechanism is insensitive to cell length, as predicted by the model.

Similarly to what was observed for wild-type cells, there was no memory with respect to elongation for $\Delta minC$ cells, as ΔL between mothers and daughters remained uncorrelated (Figure 7B). Also, as shown in Figure 7C, constant elongation in $\Delta minC$ cells was achieved by modulating the interdivision time T (and not the growth rate, data not shown). The $\Delta minC$ data showcased the striking agreement between the average cell behavior and the analytical expression of the constant extension model over a wide range of cell lengths (binned data versus black line, Figure 7C).

The $\Delta minC$ data also demonstrated that cell elongation is independent of where division occurs (Figure 7D). Even when a

division occurred at a pole (producing a minicell), the viable offspring, which inherited all of the genetic material and most of the cytoplasmic and membrane content of the mother, implemented a constant elongation just like cells generated by a normal division. Thus, it is the process of division itself, and not the partitioning of cellular content, that resets the constant extension mechanism.

Another interesting aspect of the $\Delta minC$ mutant is that it has a moderate chromosome segregation defect (Jaffé et al., 1988) and a skewed distribution of ΔL toward higher values (Figure 7E). We envision two possibilities by which a partial DNA segregation defect can result in tailed ΔL distribution.

DNA segregation may be part of a fail-safe or checkpoint mechanism that is superimposed over the constant extension mechanism. In this scenario, chromosome segregation would normally occur within the time needed to grow the appropriate ΔL and divide, having no impact on the constant extension mechanism. However, if DNA segregation becomes abnormal and does not complete within this time period, a fail-safe mechanism would override the constant extension mechanism by delaying division. Such override may occur through so-called “nucleoid occlusion” (Wu and Errington, 2012), which is known to interfere with FtsZ ring assembly in *E. coli*. A delay in division would result in higher ΔL values than expected. Note that any fail-safe mechanism that blocks a step required for division as a response to a defect would override the constant extension mechanism until the defect is resolved. A good example is the SOS response to DNA damage that blocks division until DNA repair is completed (Huisman and D’Ari, 1981).

Alternatively, DNA segregation may be an inherent part of the constant extension mechanism, with the nucleoid acting as a molecular ruler. For example, cell extension may be involved in nucleoid separation, which in turn may trigger cell constriction, perhaps by relieving some form of DNA occlusion. A defect in DNA segregation would then delay division, leading to higher ΔL values. Future studies will be required to distinguish between these two possibilities.

Defining Features of the Constant Extension Mechanism

A constant extension mechanism strongly departs conceptually from the deeply rooted critical size paradigm. In all size-based models, cells “sense” how big they are, whereas in the constant extension model, cells are blind to their size and instead “sense” how much they have grown. It is important to note that, although we are measuring cell length, cells may be “measuring” a difference in any cell size parameter; it could be a difference in cell length, but it could also be a difference in cell mass, surface area, or volume, as cell width does not change during the cell cycle.

Our findings suggest that cells follow the simple rule of triggering division when they have elongated the targeted ΔL . This is sufficient to provide a cell size homeostasis mechanism. We have identified several defining features of this cell size homeostasis mechanism. (1) Cell size deviations are abated exponentially over generations (Figure 3J). (2) The constant extension mechanism does not need to be precise (Figure 3), with experimental $CV_{\Delta L}$ of 19%–26%. A greater variability in ΔL among

cells would increase the variability in cell length at birth but would not affect the average cell length of the population or the homeostatic capability of the mechanism. (3) The constant extension mechanism provides cell size homeostasis that is robust with respect to cell length and division placement. Our $\Delta minC$ data explicitly demonstrate the robustness of the mechanism: despite aberrant cell lengths at birth, wide imprecision of division placement, and large variation in cell elongation, $\Delta minC$ cells maintain a stable cell size distribution over time (data not shown) and from culture to culture (Figure 7F). (4) Any division, including nonproductive ones that create minicells, resets the constant extension mechanism. (5) Any imprecision in size extension at one generation has no impact on the precision at the next generation (Figures S5F and S5G). These defining features provide a strong foundation for future molecular studies and will have to be accounted for by any molecular models of the constant extension mechanism.

Several lines of evidence suggest that a constant cell extension might be an ancient and broadly applicable means of achieving cell size homeostasis. *E. coli* and *C. crescentus* are evolutionary distant, having diverged more than one billion years ago. Their divisions (symmetric versus asymmetric) are distinct. *E. coli* changes its length according to nutrient availability, whereas *C. crescentus* does not. *E. coli* can undergo overlapping rounds of DNA replication, whereas this has never been observed for *C. crescentus*. Despite these profound differences in growth, division, and replication cycle, a similar cell size control operates, suggesting that this size homeostatic model is applicable to a wide variety of bacterial species under different growth conditions.

EXPERIMENTAL PROCEDURES

Strains and Media

C. crescentus CB15N (Evinger and Agabian, 1977) and *Escherichia coli* K12 strain BW25113 (Datsenko and Wanner, 2000) were used for the experiments unless indicated. *E. coli* BW25113 and $\Delta minC$, $\Delta ptsH$, and $\Delta ptsI$ derivatives were obtained from the Yale *E. coli* Genetic stock center. *C. crescentus* CB15N was grown either in PYE medium (2 g/l bacto-peptone, 1 g/l yeast extract, 1 mM $MgSO_4$, 0.5 mM $CaCl_2$) or M2G medium (0.87 g/l Na_2HPO_4 , 0.54 g/l KH_2PO_4 , 0.50 g/l NH_4Cl , 0.2% (w/v) glucose, 0.5 mM $MgSO_4$, 0.5 mM $CaCl_2$, 0.01 mM $FeSO_4$). *E. coli* BW25113 was grown in LB medium (10 g/l NaCl, 5 g/l yeast extract, 10 g/l tryptone), or M9-supplemented medium (6 g/l $Na_2HPO_4 \cdot 7H_2O$, 3 g/l KH_2PO_4 , 0.5 g/l NaCl, 1 g NH_4Cl , 2 mM $MgSO_4$, 1 μ g/l thiamine supplemented with 0.1% casamino acids and 0.2% glucose).

Microscopy

C. crescentus cells were grown up to exponential phase ($OD_{600nm} < 0.3$) and were spotted on 0.3% agarose pads containing M2G medium unless specified otherwise. Microscopy was performed on an Eclipse 80i microscope (Nikon) equipped with a phase-contrast objective Plan ApoChromat 100 \times /1.40 NA (Nikon), an Orca-II-ER (Hamamatsu Photonics), and an Andor iXon DU-897E camera (Andor Technology) with 2 \times optivar. Images were acquired every 2.5 min using MetaMorph software (Molecular Devices).

For still images of *E. coli* strains, cells were grown at 30°C up to exponential phase ($OD_{600nm} < 0.3$) and were spotted on 1% agarose pads. For microfluidic experiments, *E. coli* cells were loaded and grown for at least five generations in the microfluidic device prior to imaging. Microscopy was performed on an Eclipse Ti-E microscope (Nikon) equipped with Perfect Focus System (Nikon) and an Orca-R2 camera (Hamamatsu Photonics) and a phase-contrast objective Plan ApoChromat 100 \times /1.45 NA (Nikon). Time-lapse images were acquired every 5 s using NIS-Element Ar software (Nikon Instruments).

Stochastic Simulations

Numerical simulations of evolving cell populations were done in MATLAB using probability density distributions matching experimentally measured distributions (Figure S7 and Table S1) as described in the Extended Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Results, Extended Experimental Procedures, seven figures, one table, and three movies and can be found with this article at <http://dx.doi.org/10.1016/j.cell.2014.11.022>.

AUTHOR CONTRIBUTIONS

M.C., I.V.S., and S.K. contributed equally to this work. S.K., M.C., and C.J.-W. conceived the project, designed the experiments, and analyzed the data. M.C., S.K., and S.E.E. performed the experiments. I.V.S. and B.B. developed mathematical models, and I.V.S., B.B., and M.C. performed simulations. A.P. assisted in data analysis. M.C., I.V.S., and C.J.-W. wrote the paper. All authors commented on and approved the paper.

ACKNOWLEDGMENTS

We would like to thank Dr. Johan Elf for sharing the design of his microfluidic device, Matts Walden and Jason Hocking for help setting up the microfluidic system in our laboratory, and the Yale *E. coli* genetic stock center for the *E. coli* BW25113 derived strains. This work was, in part, supported by the National Institute of Health (R01 GM065835 to C.J.-W.). S.K. and S.E.E. were, in part, supported by the Astellas Foundation for Research on Metabolic Disorders and a predoctoral training grant of the National Institute of Health (5 T32 GM 7223-39), respectively. We are also grateful to the staffs of the Yale University Faculty of Arts and Sciences High Performance Computing Center for support and to the members of the Jacobs-Wagner laboratory for fruitful discussions and for critical reading of this manuscript. C.J.-W. is an investigator of the Howard Hughes Medical Institute.

Received: August 13, 2014

Revised: October 20, 2014

Accepted: November 13, 2014

Published: December 4, 2014

REFERENCES

- Adler, H.I., Fisher, W.D., Cohen, A., and Hardigree, A.A. (1967). Miniature *Escherichia coli* cells deficient in DNA. *Proc. Natl. Acad. Sci. USA* 57, 321–326.
- Akerlund, T., Nordström, K., and Bernander, R. (1995). Analysis of cell size and DNA content in exponentially growing and stationary-phase batch cultures of *Escherichia coli*. *J. Bacteriol.* 177, 6791–6797.
- Amir, A. (2014). Cell Size Regulation in Bacteria. *Phys. Rev. Lett.* 112, 208102.
- Bates, D., and Kleckner, N. (2005). Chromosome and replisome dynamics in *E. coli*: loss of sister cohesion triggers global chromosome movement and mediates chromosome segregation. *Cell* 121, 899–911.
- Boye, E., and Nordström, K. (2003). Coupling the cell cycle to cell growth. *EMBO Rep.* 4, 757–760.
- Cooper, S. (1991). Bacterial Growth and Division: Biochemistry and Regulation of Prokaryotic and Eukaryotic Division Cycles, First Edition, *Volume 1* (San Diego, CA: Academic Press).
- Cooper, S., and Helmstetter, C.E. (1968). Chromosome replication and the division cycle of *Escherichia coli* B/r. *J. Mol. Biol.* 31, 519–540.
- Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* 97, 6640–6645.
- Di Talia, S., Skotheim, J.M., Bean, J.M., Siggia, E.D., and Cross, F.R. (2007). The effects of molecular noise and size control on variability in the budding yeast cell cycle. *Nature* 448, 947–951.
- Donachie, W.D. (1968). Relationship between cell size and time of initiation of DNA replication. *Nature* 219, 1077–1079.
- Evinger, M., and Agabian, N. (1977). Envelope-associated nucleoid from *Caulobacter crescentus* stalked and swarmer cells. *J. Bacteriol.* 132, 294–301.
- Evinger, M., and Agabian, N. (1979). *Caulobacter crescentus* nucleoid: analysis of sedimentation behavior and protein composition during the cell cycle. *Proc. Natl. Acad. Sci. USA* 76, 175–178.
- Fantes, P.A. (1977). Control of cell size and cycle time in *Schizosaccharomyces pombe*. *J. Cell Sci.* 24, 51–67.
- Fantes, P.A., and Nurse, P. (1981). Division timing: controls, models and mechanisms. In *The cell cycle*, P.C.L. John, ed. (New York City: Cambridge University Press), pp. 11–34.
- Harris, L.K., Dye, N.A., and Theriot, J.A. (2014). A *Caulobacter* MreB mutant with irregular cell shape exhibits compensatory widening to maintain a preferred surface area to volume ratio. *Mol. Microbiol.* <http://dx.doi.org/10.1111/mmi.12811>.
- Hosoda, K., Matsuura, T., Suzuki, H., and Yomo, T. (2011). Origin of lognormal-like distributions with a common width in a growth and division process. *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys.* 83, 031118.
- Huisman, O., and D'Ari, R. (1981). An inducible DNA replication-cell division coupling mechanism in *E. coli*. *Nature* 290, 797–799.
- Jaffé, A., D'Ari, R., and Hiraga, S. (1988). Minicell-forming mutants of *Escherichia coli*: production of minicells and anucleate rods. *J. Bacteriol.* 170, 3094–3101.
- Johnston, G.C., Pringle, J.R., and Hartwell, L.H. (1977). Coordination of growth with cell division in the yeast *Saccharomyces cerevisiae*. *Exp. Cell Res.* 105, 79–98.
- Koch, A.L., and Schaechter, M. (1962). A model for statistics of the cell division process. *J. Gen. Microbiol.* 29, 435–454.
- Koppes, L.J., Meyer, M., Oonk, H.B., de Jong, M.A., and Nanninga, N. (1980). Correlation between size and age at different events in the cell division cycle of *Escherichia coli*. *J. Bacteriol.* 143, 1241–1252.
- Lord, P.G., and Wheals, A.E. (1981). Variability in individual cell cycles of *Saccharomyces cerevisiae*. *J. Cell Sci.* 50, 361–376.
- McAdams, H.H., and Shapiro, L. (2009). System-level design of bacterial cell cycle control. *FEBS Lett.* 583, 3984–3991.
- Osella, M., Nugent, E., and Cosentino Lagomarsino, M. (2014). Concerted control of *Escherichia coli* cell division. *Proc. Natl. Acad. Sci. USA* 111, 3431–3435.
- Robert, L., Hoffmann, M., Krell, N., Aymerich, S., Robert, J., and Doumic, M. (2014). Division in *Escherichia coli* is triggered by a size-sensing rather than a timing mechanism. *BMC Biol.* 12, 17.
- Santi, I., Dhar, N., Bousbaine, D., Wakamoto, Y., and McKinney, J.D. (2013). Single-cell dynamics of the chromosome replication and cell division cycles in mycobacteria. *Nat. Commun.* 4, 2470.
- Schaechter, M., Maaloe, O., and Kjeldgaard, N.O. (1958). Dependency on medium and temperature of cell size and chemical composition during balanced growth of *Salmonella typhimurium*. *J. Gen. Microbiol.* 19, 592–606.
- Schaechter, M., Williamson, J.P., Hood, J.R., Jr., and Koch, A.L. (1962). Growth, cell and nuclear divisions in some bacteria. *J. Gen. Microbiol.* 29, 421–434.
- Siegal-Gaskins, D., and Crosson, S. (2008). Tightly regulated and heritable division control in single bacterial cells. *Biophys. J.* 95, 2063–2072.
- Sliusarenko, O., Heinritz, J., Emonet, T., and Jacobs-Wagner, C. (2011). High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. *Mol. Microbiol.* 80, 612–627.
- Svecizer, A., Novak, B., and Mitchison, J.M. (1996). The size control of fission yeast revisited. *J. Cell Sci.* 109, 2947–2957.
- Terrana, B., and Newton, A. (1975). Pattern of unequal cell division and development in *Caulobacter crescentus*. *Dev. Biol.* 44, 380–385.

- Tsukanov, R., Reshes, G., Carmon, G., Fischer-Friedrich, E., Gov, N.S., Fishov, I., and Feingold, M. (2011). Timing of Z-ring localization in *Escherichia coli*. *Phys. Biol.* **8**, 066003.
- Turner, J.J., Ewald, J.C., and Skotheim, J.M. (2012). Cell size control in yeast. *Curr. Biol.* **22**, R350–R359.
- Ullman, G., Wallden, M., Marklund, E.G., Mahmutovic, A., Razinkov, I., and Elf, J. (2013). High-throughput gene expression analysis at the level of single proteins using a microfluidic turbidostat and automated cell tracking. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **368**, 20120025.
- Voorn, W.J., and Koppes, L.J. (1998). Skew or third moment of bacterial generation times. *Arch. Microbiol.* **169**, 43–51.
- Wakamoto, Y., Ramsden, J., and Yasuda, K. (2005). Single-cell growth and division dynamics showing epigenetic correlations. *Analyst (Lond.)* **130**, 311–317.
- Wang, P., Robert, L., Pelletier, J., Dang, W.L., Taddei, F., Wright, A., and Jun, S. (2010). Robust growth of *Escherichia coli*. *Curr. Biol.* **20**, 1099–1103.
- Wold, S., Skarstad, K., Steen, H.B., Stokke, T., and Boye, E. (1994). The initiation mass for DNA replication in *Escherichia coli* K-12 is dependent on growth rate. *EMBO J.* **13**, 2097–2102.
- Wu, L.J., and Errington, J. (2012). Nucleoid occlusion and bacterial cell division. *Nat. Rev. Microbiol.* **10**, 8–12.

The Structure and Regulation of Human Muscle α -Actinin

Euripedes de Almeida Ribeiro, Jr.,^{1,10} Nikos Pinotsis,^{1,10} Andrea Ghisleni,^{2,10} Anita Salmazo,¹ Petr V. Konarev,³ Julius Kostan,¹ Björn Sjöblom,¹ Claudia Schreiner,¹ Anton A. Polyansky,^{1,8} Eirini A. Gkoukoulia,¹ Mark R. Holt,² Finn L. Aachmann,⁴ Bojan Žagrović,¹ Enrica Bordignon,^{5,9} Katharina F. Pirker,⁶ Dmitri I. Svergun,³ Mathias Gautel,^{2,*} and Kristina Djinović-Carugo^{1,7,*}

¹Department of Structural and Computational Biology, Max F. Perutz Laboratories, University of Vienna, Campus Vienna Biocenter 5, 1030 Vienna, Austria

²British Heart Foundation Centre of Research Excellence, Randall Division for Cell and Molecular Biophysics and Cardiovascular Division, King's College London, London SE1 1UL, UK

³European Molecular Biology Laboratory, Deutsches Elektronen-Synchrotron, Notkestrasse 85, 22603 Hamburg, Germany

⁴Department of Biotechnology, Norwegian University of Science and Technology, Sem Sælands vei 6/8, 7491 Trondheim, Norway

⁵Laboratory of Physical Chemistry, ETH Zurich, Vladimir-Prelog-Weg 2, 8093 Zurich, Switzerland

⁶Division of Biochemistry, Department of Chemistry, University of Natural Resources and Life Sciences, Muthgasse 18, 1190 Vienna, Austria

⁷Department of Biochemistry, Faculty of Chemistry and Chemical Technology, University of Ljubljana, Aškerčeva 5, 1000 Ljubljana, Slovenia

⁸M.M. Shemyakin and Yu.A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow 117997, Russia

⁹Fachbereich Physik, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany

¹⁰Co-first author

*Correspondence: mathias.gautel@kcl.ac.uk (M.G.), kristina.djinovic@univie.ac.at (K.D.-C.)

<http://dx.doi.org/10.1016/j.cell.2014.10.056>

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

SUMMARY

The spectrin superfamily of proteins plays key roles in assembling the actin cytoskeleton in various cell types, crosslinks actin filaments, and acts as scaffolds for the assembly of large protein complexes involved in structural integrity and mechanosensation, as well as cell signaling. α -actinins in particular are the major actin crosslinkers in muscle Z-disks, focal adhesions, and actin stress fibers. We report a complete high-resolution structure of the 200 kDa α -actinin-2 dimer from striated muscle and explore its functional implications on the biochemical and cellular level. The structure provides insight into the phosphoinositide-based mechanism controlling its interaction with sarcomeric proteins such as titin, lays a foundation for studying the impact of pathogenic mutations at molecular resolution, and is likely to be broadly relevant for the regulation of spectrin-like proteins.

INTRODUCTION

Mobility is essential to all living organisms, from organelle transport to movement of entire organisms. In many motile systems, actin and myosin filaments assume ordered arrays organized by specific actin or myosin ligands. In higher animals, movement is performed by striated muscle, defined by highly regular arrangements of visible striations. The minimal contractile unit of striated muscle is the sarcomere, which is anchored and stabilized by transverse crosslinking structures at the two lateral

Z-disk boundaries, the A-band and the central M-band (Gautel, 2011; Tskhovrebova and Trinick, 2010). In vertebrates, the giant protein titin (connectin) spans Z-disks to M-bands and may act as a blueprint for sarcomere assembly (Gautel, 2011; Tskhovrebova and Trinick, 2010). Within the vertebrate Z-disk, a complicated network of protein-protein interactions anchors and stabilizes the actin and the elastic titin filaments (Luther, 2009).

α -actinin was originally described as an actin-crosslinking Z-disk protein in muscle (Masaki et al., 1967), but its four closely related isogenes (*ACTN1–4*) fulfil similar functions in all cell types (Foley and Young, 2014; Sjöblom et al., 2008). α -actinin, in particular isoform 2 (encoded by *ACTN2*), is the major Z-disk component, where it plays a central role crosslinking actin and titin filaments. α -actinin is an antiparallel homodimer of more than 200 kDa, comprising an N-terminal actin-binding domain (ABD), a central domain of four spectrin-like repeats (SRs), and a C-terminal calmodulin-like domain (CAMD) with two pairs of EF hand motifs (EFs) (Figure 1A). Because the SR region appears to have a cylindrical shape, it is also called the rod domain.

The elementary structure of the Z-disk is that of a tetragonal array of antiparallel actin filaments spaced 240 Å apart and crosslinked by successive layers of filaments at intervals of ≈ 190 Å rotated by 90° between each layer along the myofibril axis (Goldstein et al., 1979). These filaments correlate with α -actinin crosslinks, but the molecular layout of α -actinin that allows strict alternating crosslinks between actin filaments remains elusive (reviewed in Luther, 2009).

In striated muscle, α -actinin also binds differentially spliced titin Z-repeats, possibly regulating the number of crosslinking α -actinin molecules (Gautel et al., 1996). These titin Z-repeats contain a short, hydrophobic α -actinin-binding motif, which interacts with the CAMD (Atkinson et al., 2001; Sorimachi et al., 1997; Young et al., 1998). Additionally, α -actinin binds a plethora

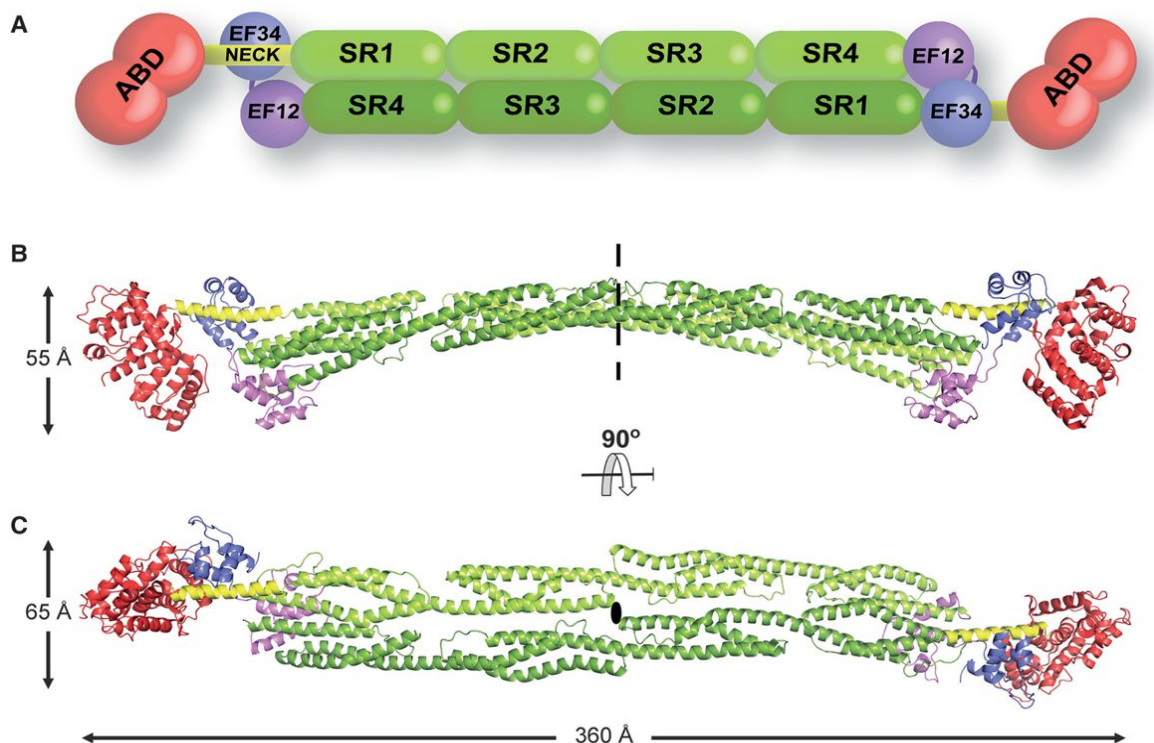


Figure 1. Complete Structure of α -Actinin-2 in Closed Conformation

(A) Domain composition of the α -actinin dimer. Color code, as in all the following figures: ABD, red; neck, yellow; SR1–SR4, green; EF1–2, violet; EF3–4, blue. (B) The dimeric structure of α -actinin-2 assembled from two halves of the α -actinin-2 protomer (ABD–SR1–SR2/SR3–SR4–CaM) through a crystallographic 2-fold axis (dashed line; ellipse in C). Overall dimensions are indicated. (C) Same as in (B), rotated 90° around the horizontal axis. See also Table S1.

of cytoplasmic and membrane proteins in striated muscle and nonmuscle tissues (Djinovic-Carugo et al., 2002; Foley and Young, 2014). To achieve ordered cytoskeletal assemblies, the binding properties of α -actinin must be spatiotemporally regulated. Actin binding of nonmuscle isoforms is regulated by binding of Ca^{2+} ions to the CAMD (Foley and Young, 2014). In contrast, muscle α -actinin is calcium insensitive, and its F-actin- and titin-binding properties are likely regulated by phospholipids (most notably phosphatidylinositol bisphosphate; PIP2) (Fukami et al., 1992; Young et al., 1998; Young and Gautel, 2000). Despite being a major integrator of titin and actin in one of the stiffest structures of the sarcomere, muscle α -actinin shows surprisingly dynamic association with the Z-disk actin cytoskeleton (Sanger and Sanger, 2008), suggesting that its actin and titin binding activity must be dynamically regulated.

Biochemical analysis led us to propose previously that the α -actinin-titin interaction is regulated by an intramolecular mechanism where a short sequence in α -actinin between the ABD and the rod interacts with the CAMD in a pseudoligand mechanism (Young and Gautel, 2000). A similar mode of interaction has been found for the α -actinin ligand palladin (Beck et al., 2011).

Here, we report the crystal structure of human α -actinin-2 at 3.5 Å resolution. It is a complete high-resolution α -actinin-2 structure, revealing insight into the mechanism that promotes

the molecular assembly of the Z-disk and the intramolecular contacts that regulate these interactions. Furthermore, the structure provides a template for the α -actinin and spectrin superfamily and insight into the impact of disease-associated genetic variants in ACTN genes.

RESULTS

Closed Structure of α -Actinin-2 Overall Architecture

The structure of α -actinin-2 was solved and refined to 3.5 Å resolution to an $R_{\text{work}}/R_{\text{free}}$ of 20.5%/25.8% (Table S1 available online). The α -actinin-2 dimer reveals a cylindrical shape ~ 360 Å long and ~ 60 Å wide (Figures 1B and 1C). Each protomer comprises an N-terminal ABD followed by an α -helical linker (neck), four spectrin-like repeats (SR1–4), and a C-terminal CAMD of two pairs of EF hands (EF1–2 and EF3–4). The first 34 and last 2 residues are missing from our model. Two antiparallel SR1–4s that assemble the core of the extended structure form the central portion of the dimer (rod). The two ABDs and two CAMDs flank the elongated assembly at its ends.

As expected, in the absence of actin, the ABD is in a closed conformation, in which the two calponin homology domains (CH1 and CH2) are in extensive contact, similar to the

arrangement found in α -actinin-3, plectin, and fimbrin ABD domains (Franzot et al., 2005; Klein et al., 2004).

The ABD is linked to the first spectrin-like repeat (SR1) of the rod through the six-turn α -helical neck that is flanked by two hinges, the first on residue G258 linking the ABD to the neck and the second on residue M283 linking the neck to SR1 (Figure 2A).

Comparison of the spectrin-like repeats in the full-length α -actinin-2 structure with the previously determined dimeric rod domain (Ylännä et al., 2001) (Protein Data Bank [PDB] ID code 1HC1) reveals high similarity between SR2, SR3, and SR4 and minor differences for SR1 (Table S2). However, although the α helices of the two SR4 domains are well aligned, there are large deviations between α helices h1 and h2 due to contacts of EF1-2 positioned over this loop in the full-length structure (Figures 1 and 2B; Figure S1A). The antiparallel α -actinin-2 dimer assembles predominately via the rod domain. The two ABDs are flanked at both termini along the long axis of the rod and are stabilized in their position by a few polar interactions with the neck region and EF3-4.

The C-terminal CAMD is divided into structurally distinct N- and C-terminal lobes: EF hands 1-2 and 3-4 connected by a short linker. The overall structure of each lobe is well defined through the main-chain atoms on the α helices of the EF hands. As expected, no density conforming to bound calcium was detected, and the architecture of the CAMD resembles calcium-free conformations of calmodulin (see below).

Calmodulin-like Domain Conformation and Interactions

The EF3-4 accommodates the α helix of the neck from the juxtaposed α -actinin-2 protomer through its hydrophobic cleft. This classifies the EF3-4-neck interaction as 1:1 or canonical (Hoeftlich and Ikura, 2002), involving the cavity formed by the first helix of the EF3-4, the linker between the EF3-4 helices, and the C-terminal helix of the lobe (Figures 2C and 2D). The EF3-4-neck interface is mainly hydrophobic, supported with a few H-bonds (Figure 2C). The neck region displays the known hydrophobic Ca^{2+} /calmodulin (CaM)-binding motif, termed 1-4-5-8 (Bayley et al., 1996), with hydrophobic A266, I269, C270, and L273 at these positions (Figure 2F). Additionally, EF3-4 interact with the connecting loop of α helices 2 and 3 from SR1 via an H-bond and a π -stacking interaction, supporting previous reports that SR1 stabilizes the interaction with EF3-4 (Young and Gautel, 2000). The overall binding interface of EF3-4 is $\sim 500 \text{ \AA}^2$, representing 11% of its total surface area. The root-mean-square deviation (rmsd) with the C-terminal lobe in complex with Zr-7 of titin (Atkinson et al., 2001) (PDB ID code 1H8B) is 1.1 \AA for 64 equivalent $\text{C}\alpha$ atoms, implying no significant conformational difference.

EF1-2 are less well defined in the electron density compared to the C-terminal lobe, and similar to EF3-4 show a binding interface of $\sim 500 \text{ \AA}^2$ with SR4. The interactions occur between the N-terminal α helix of EF1-2, the loop connecting EF1 and EF2 intercalates between α helices 2 and 3 of SR4, whereas the C-terminal α helix of EF1-2 lies in a parallel orientation on the third α helix 3 of SR4 (Figure 2B). In essence, only the N- and C-terminal helices of EF1-2 are involved in the binding, and this mode has no resemblance to any known classification of calmodulin-like domains but rather is similar to a bound-free lobe (Hoeftlich and Ikura, 2002).

Interestingly, even though EF1-2 and EF3-4 have different interaction interfaces, they both adopt the same semiopen conformation (Chin and Means, 2000; Swindells and Ikura, 1996) (Figure S1B). The interaction of EF3-4 with the neck resembles that with titin Zr-7, both in terms of domain structure and ligand binding (Atkinson et al., 2001) (Figure 2F). EF1-2 and EF3-4 are mostly similar to the C-terminal lobe of human cardiac troponin C (TnC) (rmsd 3.1 and 3.2 \AA for EF1-2 and EF3-4, respectively, with the C-terminal TnC lobe; PDB ID code 1J1E) (Takeda et al., 2003), confirming a previous observation that side-chain clusters in the EF hands are not related to the semi-open conformation (Atkinson et al., 2001). Further details on CAMD conformation can be found in Supplemental Information.

Sequence alignment of the EF hand pairs from α -actinin-2 and TnC reveals that in α -actinin-2, several calcium ligands in TnC, typically D, N, and E, are replaced by bulkier charged or smaller residues (Figure 2E). In essence, an R inhibits calcium binding in EF1 and a longer Q in EF2, where also a smaller S occupies the last position. In EF3, two positions typically occupied by a negatively charged D host an A and a P, whereas no residue promoting calcium binding is found in EF4.

Molecular Determinants of EF3-4 Ligand Specificity

Young and Gautel (2000) showed that CAMD binds to the neck region with lower affinity than to Zr-7 (K_d of $0.57 \text{ }\mu\text{M}$ and $0.19 \text{ }\mu\text{M}$, respectively). In order to understand the higher affinity of CAMD for Zr-7, we compared the interaction interfaces of CAMD with Zr-7 (Atkinson et al., 2001) and the neck region.

Structural superposition of EF3-4 domains in complex with the ligands shows a similar layout between the α helices of the neck and Zr-7 in complex with CAMD (Figure 2F). Comparison of the 1-4-5-8 motifs shows that titin Zr-7 hosts a bulkier hydrophobic residue at position 1 (V702/A266), whereas position 5 in the neck is a less bulky Cys residue at position 270 (cf V706 in Zr-7) (Carugo, 2014), resulting in reduced stabilization of the interaction. This is reflected in the increase of the probability measure $P_{\Delta G, \text{IF}}$ for the interface derived from the gain in solvation energy upon complexation, where $P_{\Delta G, \text{IF}} > 0.5$ points to hydrophilic/unspecific and $P_{\Delta G, \text{IF}} < 0.5$ points to hydrophobic/specific interfaces using PISA (Krissinel and Henrick, 2007). Calculated for the ligand, these range from 0.31 in the EF3-4-Zr-7 complex to 0.49 for the EF3-4-neck complex, indicating a reduced specificity with a concomitant increased hydrophilic nature of the interaction.

PIP2 Binding Site

The model of α -actinin-2 and -3 activation by PIP2 hypothesizes that PIP2 docks with the polar head group on the CH2 domain, whereas its aliphatic chain reaches the CAMD binding site on the neck, perturbing this interaction (Franzot et al., 2005). This regulatory mechanism requires spatial proximity of the PIP2 binding site and neck region. The PIP2 binding site was originally mapped to residues 165–181 (Fraley et al., 2003; Fukami et al., 1996) on a loop connecting the first and second α helix of the CH2 domain. Structural analysis of the ABD of α -actinin-3 suggested a triplet of positively charged residues, which form a platform for PIP2 binding (α -actinin-2 residues R163, R169, and R192) (Franzot et al., 2005).

We used molecular dynamics (MD) simulations and flexible ligand docking to place PIP2 into the structure of α -actinin-2. Although the results suggest that the PIP2 binding site is likely

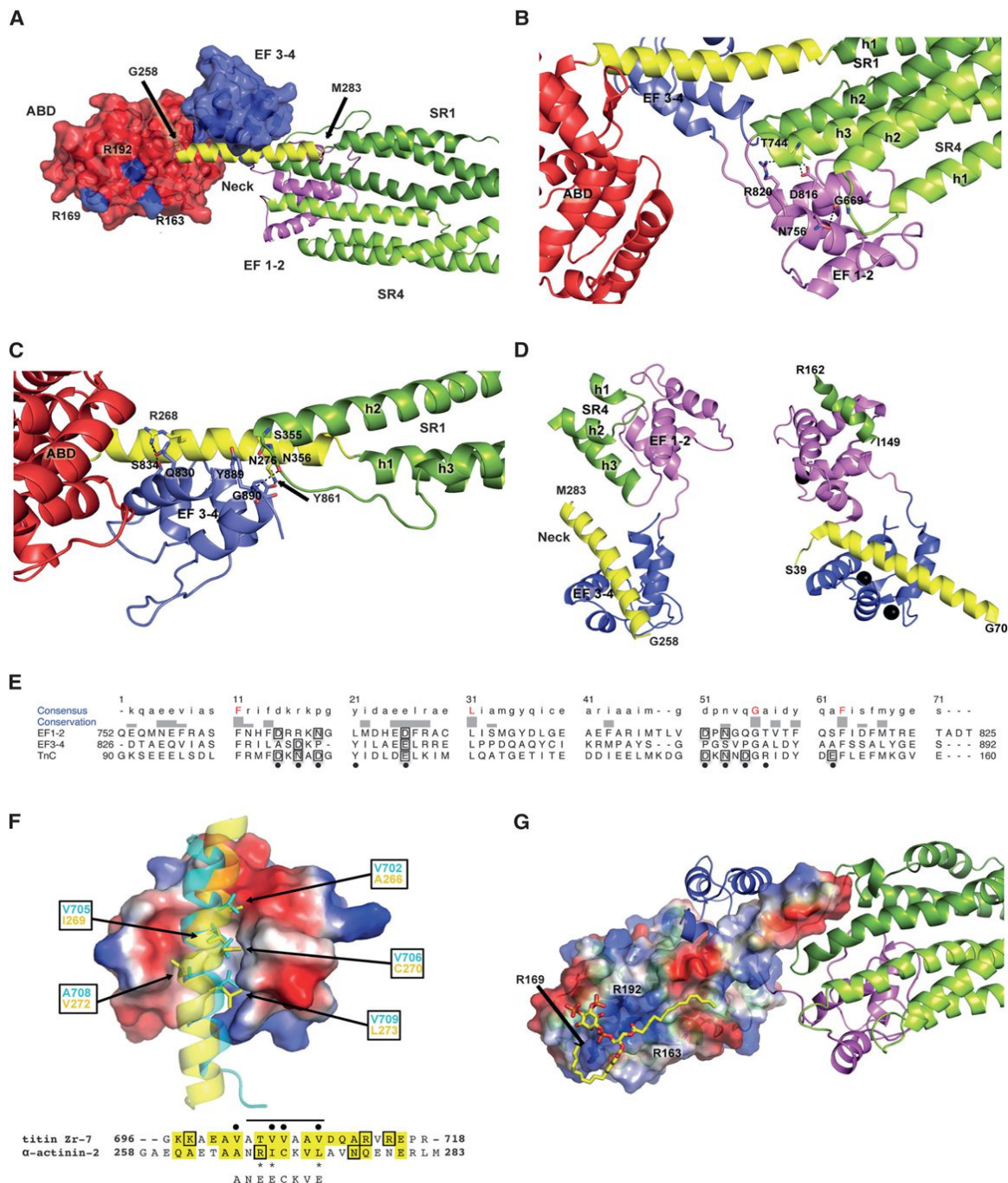


Figure 2. Close-Up of the Functional Domain Interactions

(A) PIP2 binding site on α -actinin-2 and the EF3-4-neck interaction. ABD and EF3-4 are presented with their solvent-accessible surface areas. The R residues responsible for PIP2 binding are highlighted in blue on the ABD surface.

(B) Detail of EF1-2 interactions with SR4.

(C) Detail of EF3-4 interactions with the neck region and SR1.

(D) Comparison of α -actinin-2 CAMD with TnC bound to TnI, aligned on EF3-4 and the C-terminal lobe of TnC. Left: cartoon representation of α -actinin EF1-2 and EF3-4 with the interacting portion of the neck from the juxtaposed subunit (yellow) and a part of SR4 from the same subunit. Right: cartoon representation of TnC bound to TnI. N-terminal lobe, violet, as in EF1-2, and the C-terminal lobe, blue, as in EF3-4. Bound N-terminal TnI fragment, yellow; C-terminal TnI helix, green. Calcium ions are shown as black spheres (on TnC). The C-terminal lobe of TnC is aligned to EF3-4. To show the direction of bound helices, residues defining the neck domain of actinin and TnI fragments are indicated.

(E) Sequence alignment of EF1-2 and EF3-4 and the C-terminal lobe of TnC. The corresponding calcium-binding positions in Ca^{2+} /CaM are indicated by black dots. Charged residues involved in calcium binding are boxed. Fully conserved residues are highlighted in red.

(legend continued on next page)

not rigid, a large fraction of docked poses exhibits important commonalities. Namely, in about 40% of 10,000 generated models of the complex, the polar PIP2 head directly interacts with the above-mentioned arginine platform. This was supported by fluorescence anisotropy using PIP2 binding site mutants (Figures S2A and S2B). At the same time, one (~35% of models; Figure 2G) or both (~4% of models; Figure S2F) PIP2 aliphatic chains, which span some 17 Å, lean on the partially hydrophobic surface of the ABD and extend toward the 1-4-5-8 motif in the neck region (Figure 2G). A similar binding mode, where not only the polar head is involved but also the aliphatic chain, has been observed in the matrix domain of HIV-1 (Saad et al., 2006). This suggests that the architecture of the α -actinin-2 wild-type (WT) provides a suitable spatial orientation of both PIP2 and CAMD binding sites.

Structure of Activated α -Actinin

The biochemical model suggests that α -actinin is activated by PIP2 binding to ABD, resulting in a release of EF3-4 from the neck, thus facilitating its interaction with titin (Young and Gautel, 2000).

Structure-guided mutants were designed to disrupt key contacts between the neck segment of α -actinin-2 and EF3-4, producing a constitutively open variant. In particular, one positively charged residue (R268) and two hydrophobic residues (I269, L273) were replaced by negatively charged glutamates (R268E/I269E/L273E), hereafter termed NEECK (Figure 2F). NEECK was used to validate the closed conformation observed in the crystal structure of WT α -actinin and to probe in cellula the biological relevance of the opening/closing mechanism.

Conformational Switch of α -Actinin Is Modulated by PIP2 and Titin Zr-7

We explored whether the molecular architecture observed in the crystal structure of α -actinin-2 is maintained in solution and how this is altered in NEECK as a model of the open state.

We used site-directed spin labeling and electron paramagnetic resonance (EPR) spectroscopy to obtain structural information on WT and NEECK α -actinin. All possible pairwise distances between labeled cysteines were computed (Figure 3A) and compared to those obtained by Q band double electron-electron resonance (DEER) experiments, showing a bimodal distribution peaking at 30 Å and below 20 Å (Figure 3B; Table S3). The experimental and computed distributions fit well in the distance range up to 35 Å (Figure 3B), confirming that the structures of α -actinin-2 in the crystal and solution are comparable.

We focused on the distance range below and around 20 Å, because the distance between C270 in the neck and C862 in EF3-4 is 12 ± 0.2 Å (Figure 3A). This spin label pair could sense

open and closed conformations of α -actinin-2. Other spin label sites within a 20 Å distance are buried or located on the rigid rod domain (Table S3).

An open conformation of NEECK was inferred from a decreased fraction of short distances between C270 and C862 in the DEER distribution compared with WT (Figure 3C). This change in distances was validated with low-temperature continuous-wave (cw) EPR (Figures S3A and S3B).

The analysis of the crystal structure using MD/docking suggests that the PIP2 binding site on ABD maps at a suitable position and distance from the CAMD-neck interaction to sense the hydrophobic tail of PIP2 (Figures 2A and 2G; Figures S2C–S2F). Can PIP2 alone induce opening of α -actinin-2? DEER measurements were carried out on WT α -actinin-2 using the more hydrophilic PIP2 analog Bodipy-TMR-PIP2-C16 (PIP2-C16*; Figure 3C; Figures S2 and S3). No significant changes were detectable in the short distance range (Figure 3C; Figure S3).

We tested whether titin Zr-7 could act on the conformational equilibrium of α -actinin-2. We addressed this question structurally by EPR spectroscopy and quantitatively by microscale thermophoresis (MST). Addition of Zr-7 plus PIP2-C16* to α -actinin-2 significantly reduced the short distance peaks in the DEER distance distribution, indicating conversion to an open conformation. A similar effect was observed after addition of 15-fold molar excess of Zr-7 to α -actinin alone in the absence of PIP2-C16* (Figure 3C; Figures S3C and S3D). This decrease was validated by cw EPR at low temperature (Figure S3E).

The effect of PIP2-C16* on α -actinin affinity to Zr-7 was quantified by MST (Figures 3D–3F; Table S4). PIP2-C16* bound to WT α -actinin with K_d 2.96 ± 0.26 μ M (Figure 3D). The results showed a significantly higher affinity of titin Zr-7 for the PIP2-C16*- α -actinin complex (Figure 3F; K_d 0.38 ± 0.06 μ M) compared to α -actinin alone (K_d 2.90 ± 0.12 μ M). The nanomolar Zr-7 binding affinity for the PIP2-C16*- α -actinin complex is comparable to that observed for the isolated CAMD (EF1–4) and Zr-7 (Figure 3E; K_d 0.24 ± 0.04 μ M), whereas NEECK shows an intermediate affinity (Figure 3E; K_d 0.92 ± 0.02 μ M) in the absence of PIP2-C16*. The results for CAMD agree with earlier binding studies (Table S4). MST thus confirms that PIP2-C16* increases Zr-7 affinity for α -actinin-2 ~10-fold, whereas the PIP2 mutant—having lower PIP2-C16* affinity (Figure 3D)—showed an insignificant increase of Zr-7 binding by PIP2-C16* (Figure 3F; Table S4).

We used small-angle X-ray scattering (SAXS), multiangle static laser light scattering (MALLS), and size-exclusion chromatography (SEC) on WT and NEECK α -actinin to understand the structural differences between closed and open conformations. The derived molecular parameters are given in Figures 4A and 4B and Table S5. SEC-MALLS shows that both α -actinin

(F) Comparison of interactions between Zr-7 and the neck α helices with EF3-4. Electrostatic surface representation of the EF hands and cartoon representation for Zr-7 (cyan) and the α -actinin-2 neck (yellow). Side chains of key hydrophobic residues are shown as sticks; sequence numbers are boxed. Structural sequence alignment between titin Zr-7 and the neck. Residues involved in the interface with EF3-4 are highlighted in yellow. Residues involved in H bonding are boxed. Asterisks denote the mutations in the NEECK mutant. Black dots denote the CaM-binding motif 1-4-5-8. Underneath is shown the sequence of the NEECK mutant. (G) α -actinin-2 with docked PIP2 (the overall top-scoring pose is shown as a yellow stick model) together with the EF3-4-neck interaction. The top-scoring pose with two PIP2 tails in contact with the neck region is presented in Figure S2F. ABD and the neck region are presented with their solvent-accessible surface areas colored by electrostatic potential and the rest by cartoon representation and color coded as in Figure 1. The three R residues responsible for PIP2 binding are indicated.

See also Table S2.

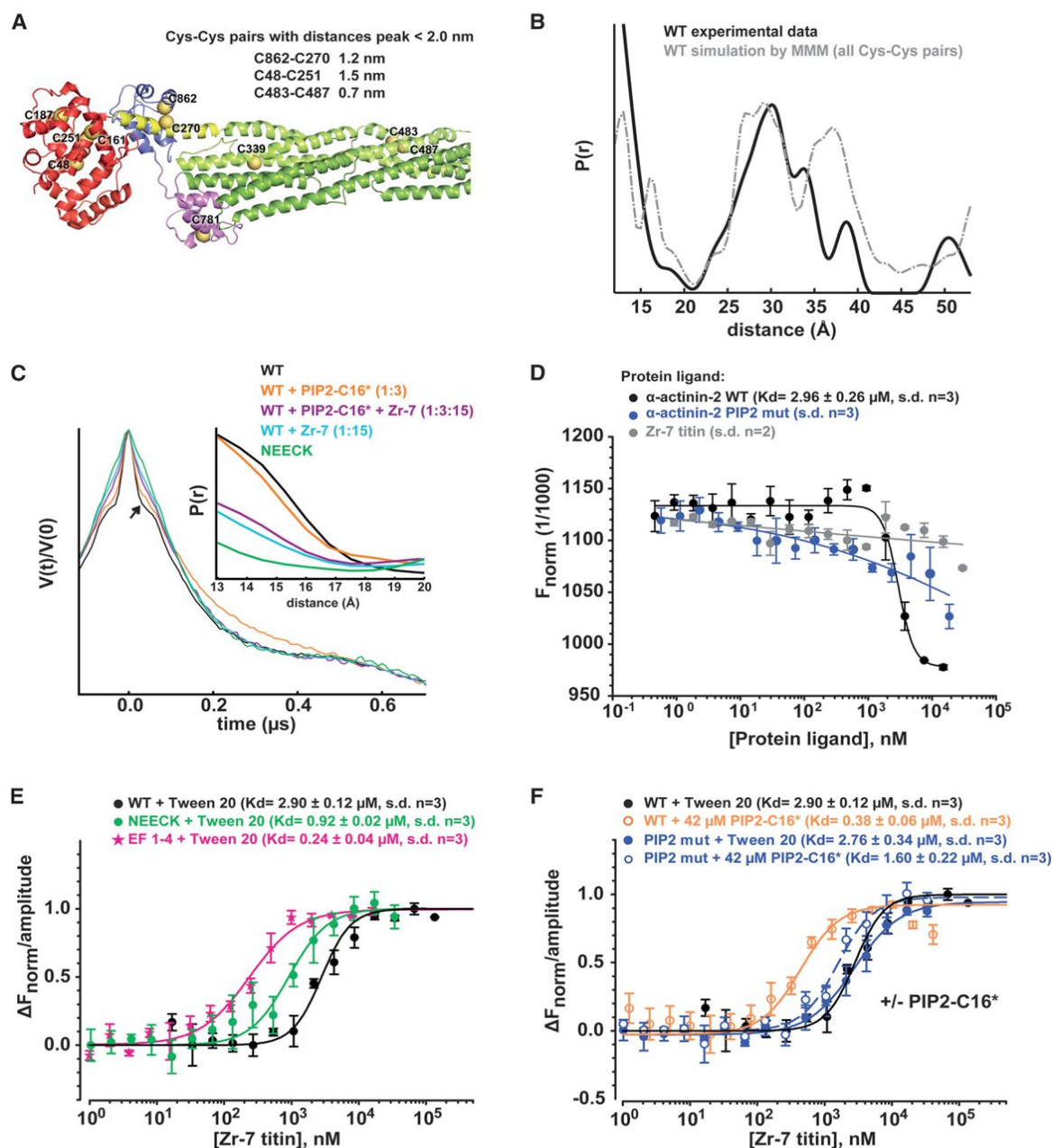


Figure 3. Structural Plasticity and Regulation of α -Actinin-2 Assessed by EPR and MST

(A) Cluster of the ten cysteine residues (Cys in dark yellow spheres) of α -actinin-2 used for the computed DEER distance distribution shown in (B). The inset shows pairs with interspin distances <20 Å.

(B) Experimental distance distribution of spin-labeled α -actinin WT (black) and simulation of the distance distribution (gray) based on spin-labeled cysteine residues from the crystal structure using the program MMM (see Extended Experimental Procedures and Table S3).

(C) DEER traces (for better comparison, adjusted by the modulation depth) and distance distributions from Q band DEER experiments using a DEER dipolar evolution time of 1 μs at 50 K of α -actinin-2 WT (black), WT plus PIP2-C16* (orange), WT plus PIP2-C16* plus Zr-7 (magenta), WT plus Zr-7 (light blue), and the NEECK mutant (green). The arrow indicates the change in the time domain trace, which is reflected in the variation of the fraction of distances <20 Å (inset).

(D) PIP2-C16* binding to α -actinin-2 measured by MST.

(E) CAMD, α -actinin-2 WT, and NEECK variant binding to Zr-7 measured by MST. The affinity determined for α -actinin-2 (+PIP2-C16*) binding to Zr-7 is in a similar range of affinity for α -actinin-2 CAMD and is implicated in Zr-7 interaction, as well as for the NEECK variant.

(F) α -actinin-2 variant binding to Zr-7 with and without PIP2-C16* measured by MST. Unlabeled Zr-7 was titrated into a fixed concentration of fluorescently labeled α -actinin-2 (50 nM).

Average and error bars (SDs) of three MST experimental replicas are plotted. Mean and SD of K_d values were calculated from these plots.

See also Tables S3 and S4.

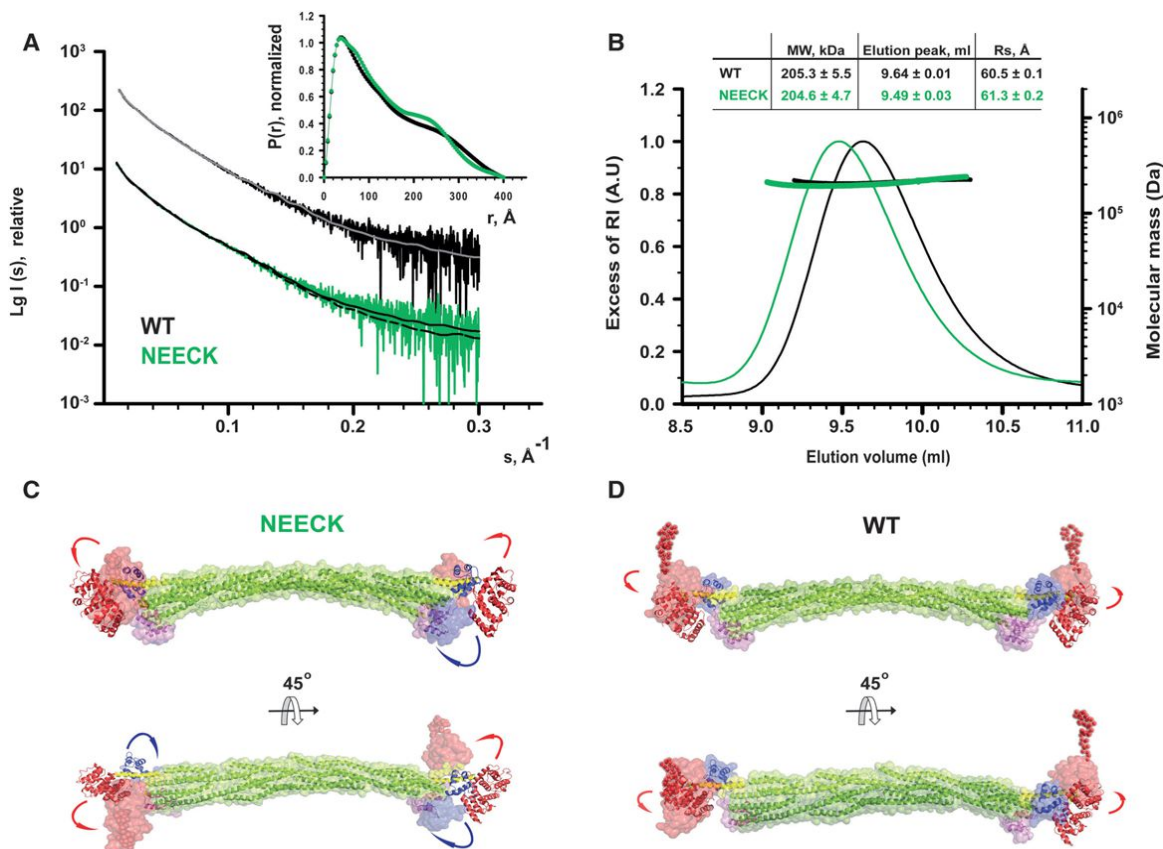


Figure 4. Solution Structure of α -Actinin-2 and the NEECK Mutant Derived from SAXS

(A) Experimental SAXS data of WT (black) and the NEECK mutant (green) of α -actinin-2. SAXS curves are computed from a rigid-body (RB) model for WT (gray) and NEECK (black). The logarithm of scattering intensity (I) is plotted as a function of the momentum transfers (s , \AA^{-1}). Successive curves are displaced by one logarithmic unit for better visualization. Distance distribution functions (inset) $P(r)$ for WT and NEECK assume slightly different shapes. RB modeling fits the experimental WT data with χ 1.25 (gray line) and experimental NEECK data with χ 1.14 (dashed black line). The fit discrepancy for NEECK increased to 1.32, assuming a helical neck (solid black line).

(B) Characterization of hydrodynamic properties of α -actinin WT and the NEECK mutant by SEC-MALLS. The lines across the protein elution volume show the molecular masses (MWs) of proteins. SEC-MALLS shows that NEECK has the same molecular weight as WT α -actinin-2 but a higher Stokes radius R_s (inset; data are represented as mean \pm SD of three experiments), corroborating the open conformation for NEECK suggested by SAXS (C). AU, arbitrary units.

(C) RB model of NEECK in solvent-accessible surface representation. The neck region was modeled as a flexible linker between the rigid bodies ABD and rod, with no contact restraint. Only one RB model is shown for clarity out of three independent BUNCH runs (Figure S5A).

(D) The best RB model of WT α -actinin-2 in solvent-accessible surface representation superimposed on the crystal structure. For WT RB modeling, only ABD was allowed a variable position, whereas EF hands 3–4 were fixed in contact with the neck.

In all models, N-terminal residues missing from the crystal structure were modeled as dummy atoms. Arrows highlight the movement of ABD and EF hands 3–4 relative to the superimposed crystal structure. See also Table S5.

samples display the same molecular mass but that NEECK has an increased Stokes radius. In order to model the NEECK variant, we first addressed the structural consequences for the α -helical neck upon release of EF3-4 by performing NMR of the WT free neck peptide (amino acids A259–Y286). Analysis of 2D nuclear Overhauser effect spectroscopy (NOESY) showed no evidence of stable secondary structure (Figure S4). NEECK was thus modeled with a flexible neck region (Figure 4C), whereas the neck for WT was modeled as a rigid-body α helix, as in the X-ray structure (Figure 4D). The best WT model fits with the crystal structure (χ = 1.25). In NEECK, the ABDs deviate from the linear alignment with the rod and EF3-4 are in open conformation (Figure 4C), giving a fit of χ = 1.14. This agrees

with increased mobility of NEECK and unfolding of the neck. In addition, we found a change in spin label mobility between WT and NEECK α -actinin using cw EPR room temperature measurements (Figure S5F). The narrower lines in the EPR spectrum of NEECK (Figure S5F) indicated increased mobility compared to WT, in line with an open conformation of NEECK. Thus, the solution structure of NEECK can best be modeled by an ensemble of conformations (Bernadó et al., 2007) where ABD and CAMD adopt various orientations (Figures S5A and S5E).

In conclusion, NEECK adopts a constitutively open conformation, and although it has to be seen as a “hyperactive” state, it is likely to approximate the Zr-7-bound structure of α -actinin in the Z-disk.

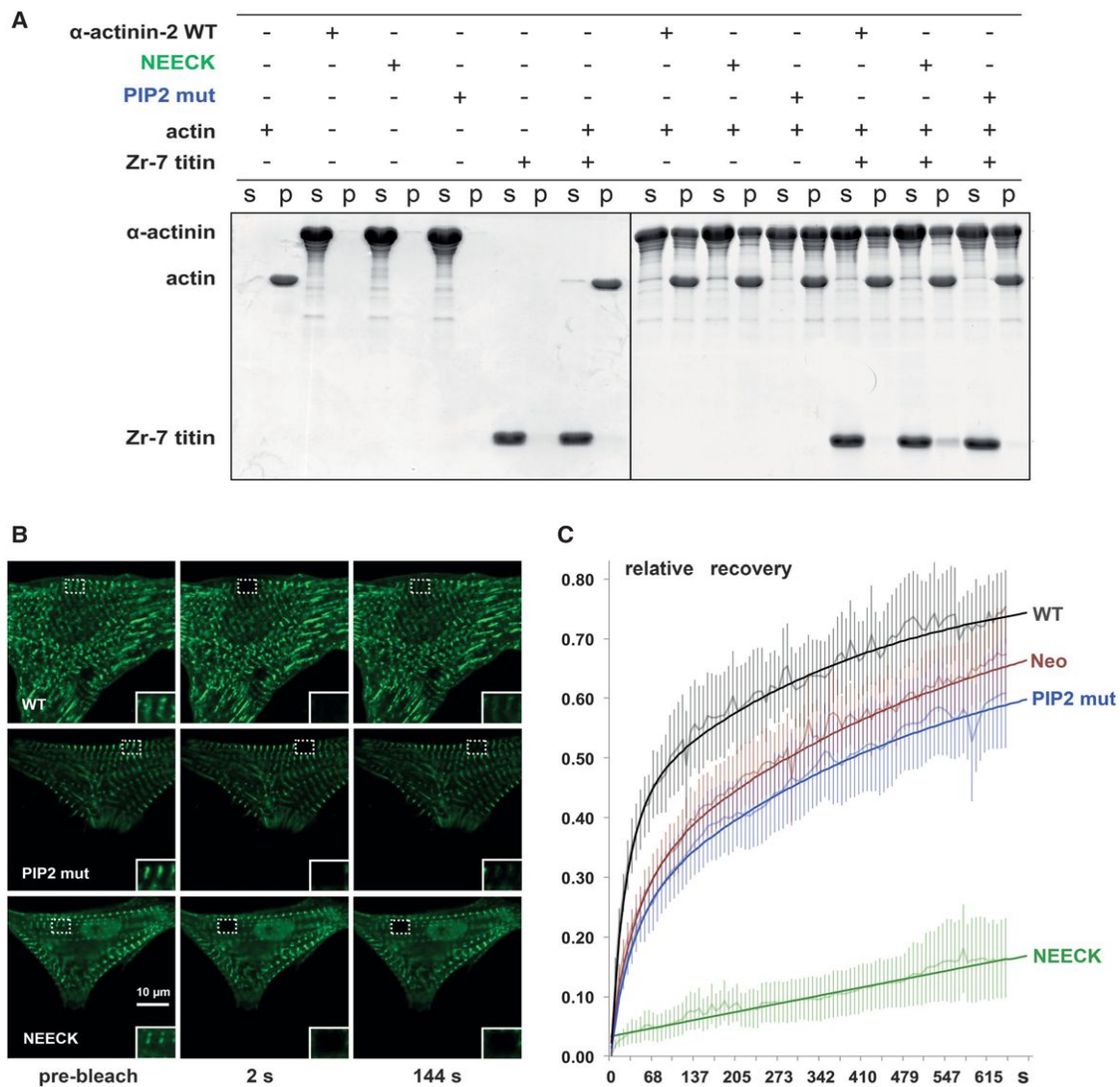


Figure 5. Mutations Affecting Regulation of α -Actinin-2 with PIP2 Do Not Influence F-Actin Binding but Impact α -Actinin-2 Z-Disk Dynamics
(A) Binding of α -actinin-2 variants to F-actin and titin Zr-7. α -actinin-2 WT, NEECK, and PIP2 mutants (PIP2 mut) were cosedimented with actin, and equal amounts of supernatant (s) and pellet (p) fractions were subjected to SDS-PAGE and visualized by Coomassie blue.
(B and C) FRAP measurements of α -actinin-2 dynamics in live NRCs expressing GFP-labeled α -actinin-2 variants (WT, PIP2 mutants, and NEECK).
(B) Snapshots at prebleach and two time points postbleach; the bleached region of interest (ROI) is highlighted by a dotted box. Note that NEECK fluorescence does not recover within the 144 s time course shown here, whereas rapid recovery is observed for WT α -actinin. Insets: ROIs enlarged 2-fold.
(C) Quantification of fluorescence intensity recovery. Note that the slowed fluorescence recovery of the PIP2 mutant is mirrored by treatment with 500 μ M neomycin (Neo). Bold lines, exponential fits; shaded lines, average values. Error bars indicate SD.

Impaired α -Actinin Regulation Disrupts Coordinated Z-Disk Assembly
To directly test the role of α -actinin regulation in Z-disk assembly, we performed live-cell imaging. We hypothesized that disrupted coordination of α -actinin-2-ligand interactions in NEECK, which binds titin constitutively, and the PIP2 mutant, where activation by PIP2 is blunted, should reduce dynamic exchange of α -actinin at Z-disks.
We assessed whether the NEECK and PIP2 mutants can still play their basic roles—binding to F-actin and titin Zr-7—by actin

cosedimentation assays. F-actin binding was unaffected in both mutants (Figure 5A), in contrast to PIP2-binding mutants in non-muscle α -actinin (Fraleigh et al., 2003). No cosedimentation of Zr-7 was seen with WT or PIP2 mutant, in contrast to NEECK. Together, these results suggest that the PIP2 or NEECK mutations do not impair F-actin binding, and that NEECK induces PIP2-independent Zr-7 binding.
We carried out fluorescence recovery after photobleaching (FRAP) studies in neonatal rat cardiomyocytes (NRCs) expressing the GFP-labeled α -actinin-2 mutants to assess the impact of

α -actinin-2 regulation on spatiotemporal dynamics in sarcomeres. The exchange of WT α -actinin-2 at Z-disks was rapid, with a fast component $t_{1/2}$ 25 ± 2 s (Figures 5B and 5C). In contrast, the NEECK mutant was dramatically slower, with $t_{1/2} > 6,134$ s. The PIP2 mutant showed reduced dynamics, with $t_{1/2}$ 35 ± 4 s (Figure 5B). The slower dynamics of α -actinin-2 mutants compared to WT observed in cellula seems not to be due to F-actin binding activity, because both mutants bind F-actin (Figure 5A). Analysis of FRAP kinetics revealed standard fast and slow components in the case of WT and the PIP2 mutant, whereas for NEECK only a slow component could be discriminated. The slow, single-exponential exchange of NEECK agrees with a dominant, high-affinity interaction. Because the NEECK interaction with titin is constitutive, the slow cellular dynamics likely reflect the high affinity of the EF3-4 interaction. Exchange of the PIP2 mutant was also slower, in agreement with reduced phospholipid regulation. To probe the role of PIP2 in regulating α -actinin dynamics independently, we used the aminoglycoside neomycin, an inhibitor of PIP2 signaling (Li and Russell, 2013; Schacht, 1976). Neomycin resulted in slower FRAP recovery of WT α -actinin-2 ($\approx 40 \pm 4$ s), similar to the PIP2 mutant (Figure 5C), supporting the notion that α -actinin-2 Z-disk dynamics are strongly dependent on PIP2 regulation, in agreement with the Z-disk localization of PIP2 (Figure S6A).

Z-disk morphology in α -actinin-transfected NRCs showed a striking phenotype for the NEECK mutant but not WT or the PIP2 mutant. Cells expressing NEECK showed gradual appearance of sarcomeres with wide α -actinin labeling, where titin epitopes peripheral (T12 antibody) and more central (Z1Z2) to the Z-disk (Young et al., 1998) were resolved as doublets flanking the edge of the α -actinin-labeled central Z-disk. This resulted in formation of actin/ α -actinin bundles resembling nemaline rods but containing diffusely localized Z-disk titin (Figures 6A–6C) and ultimately complete disruption of sarcomeres (Figure 6C). Vinculin localization in NEECK-transfected cardiomyocytes was unaffected (Figure S6B).

Whereas the optically resolvable distance between T12 epitopes in WT-transfected cells is ~ 200 nm, in NEECK-transfected cells this was > 600 nm, and > 800 nm after 3 days (Figure S6C). Similar splitting to > 200 nm was also seen for Z1Z2, normally at the limit of optical resolution with a separation of ~ 100 nm (Young et al., 1998). This suggests that the ordered integration of titin and α -actinin is severely disrupted by NEECK, raising the question of whether the spatiotemporal integration of other titin Z-disk ligands is affected. Current models of titin layout in the Z-disk predict that Z-disk widening could only be achieved by relative slipping of the overlapping N termini of titin molecules entering the Z-disk from two antiparallel sarcomere halves (Gautel, 2011). Titin molecules are crosslinked in an antiparallel palindromic complex of domains Z1-Z2 and the small Z-disk protein telethonin (Zou et al., 2006). Both telethonin and titin-Z1Z2 epitopes strictly colocalize at the Z-disk periphery (Figure 6D) but also in the wide NEECK Z-disks (Figure 6E). These results show that intramolecular autoregulation of α -actinin-ligand interactions is crucial for sarcomere integrity regulating the integration of titin, actin, and α -actinin in Z-disks of controlled width, without affecting interaction of titin with Z-disk proteins such as telethonin.

DISCUSSION

The structure of α -actinin-2 shows a modular architecture, yet is more than just “the sum of its parts”: important intra- and inter-molecular contacts lock the molecule in a closed conformation that is crucial for dynamic regulation.

Pseudoligand Model Validation

The structure of α -actinin-2 displays a closed, autoinhibited conformation, as suggested by Young and Gautel (2000) (Figure 1B). The closed structure of α -actinin-2 shows, furthermore, that the PIP2 binding site on the ABD (Franzot et al., 2005) maps at a suitable position and distance from the CAMD-neck interaction for sensing the hydrophobic tail of PIP2 (Figure 2A), as supported by MD/docking simulations (Figure 2G; Figures S2C–S2F).

The closed α -actinin-2 conformation shows that a number of interactions between SR1–4 stabilize the formation of antiparallel dimers, providing its structural rigidity and stability. Although addition of PIP2-C16* alone does not promote complete opening of α -actinin-2, at least as measurable by DEER, it promotes binding of Zr-7 with nanomolar affinity (Figure 3F), suggesting a positive allosteric modulation for opening and ligand binding. Furthermore, local structural changes in the CH2 domain, as observed upon PIP2 binding to nonmuscle α -actinin (Full et al., 2007), cannot be excluded.

Additional mechanisms might act in cells that cooperate with PIP2 or offer alternative regulation, including posttranslational modifications or protein cofactors. However, we could not identify any such plausible sites conserved between muscle α -actinin-2 and -3 in proteomic databases. No protein cofactors regulating α -actinin-titin interactions have been identified to date.

Structural comparison of interactions between EF3-4, Zr-7, and the neck reveals the basis for the higher affinity of α -actinin-2 for titin versus the pseudoligand neck (Young and Gautel, 2000) (Figure 2F). Our results suggest a model for PIP2 regulation of α -actinin, relying on structural plasticity and conformational dynamics of α -actinin-2 (Figure 7A): in the absence of PIP2, α -actinin-2 exists in two conformational states, a highly populated closed [A_c] and a low populated open and active state [A_o]. Addition of PIP2 triggers an activated state [A^* :PIP2] with lower activation energy for opening. Binding of Zr-7 recruits α -actinin-2 in the open conformation [A_o :Zr-7], which is enriched, because α -actinin-2 binds Zr-7 with higher affinity than it does to the pseudoligand neck (Figure 2F; Table S4).

Implications for Binding F-Actin Crosslinking and Z-Disk Structure

In the Z-disk, antiparallel actin filaments are crosslinked by α -actinin in a paracrystalline tetragonal lattice (Goldstein et al., 1988). To analyze the structural determinants of α -actinin's principal crosslinking function, we mapped the known actin binding sites (ABSs) to the structure of the α -actinin dimer. ABSs in ABDs of several actin-binding proteins have been located on the first and the last α helix of the CH1 domain and on the first α helix of the CH2 domain (Sjöblom et al., 2008). ABSs in α -actinin-2 are exposed, and not blocked by interdomain interactions (Figure 7B).

To generate a 3D model of F-actin/ α -actinin, we superimposed the structure of α -actinin-2 on that of the F-actin-bound

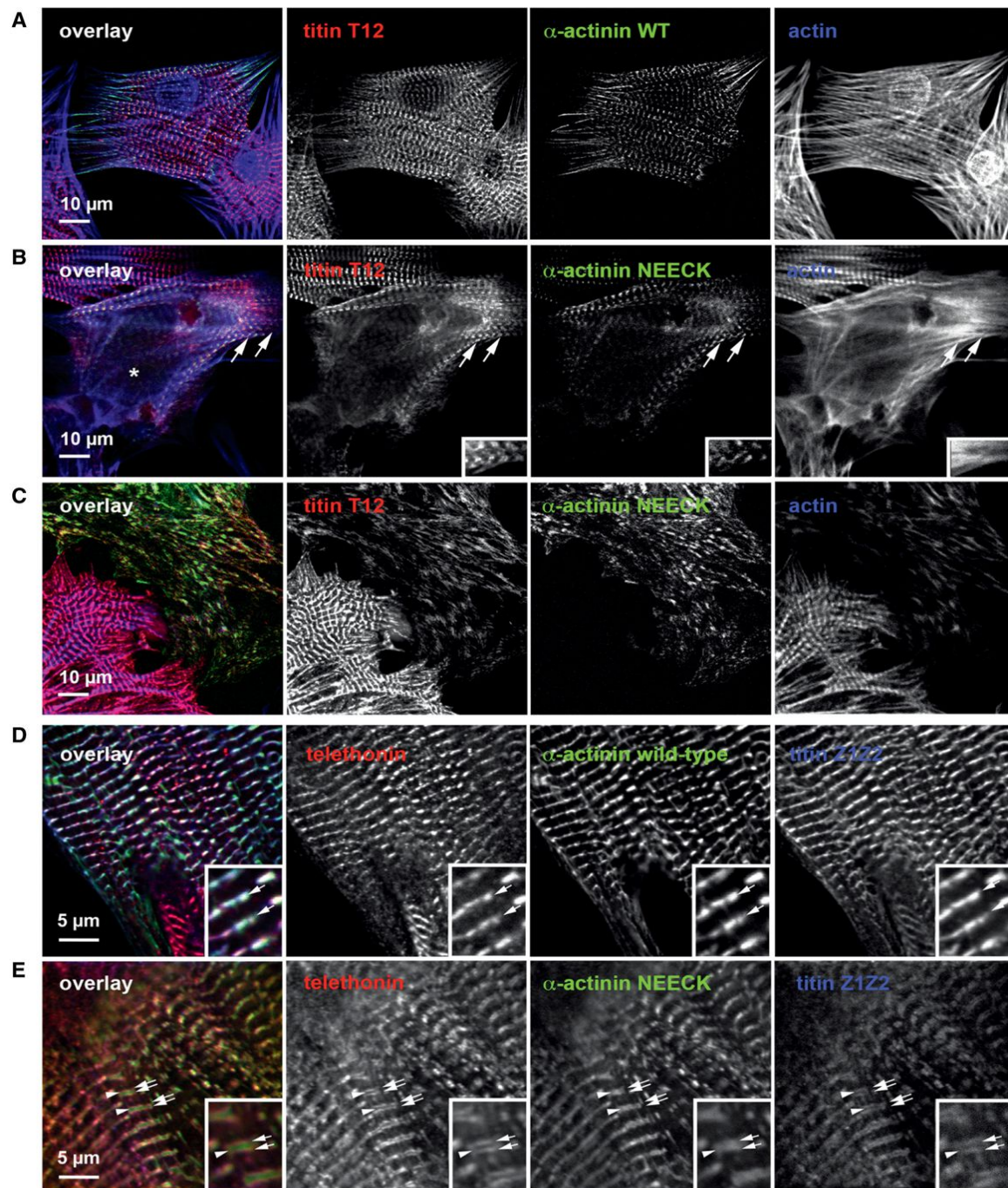


Figure 6. Constitutively Activated α -Actinin-2 Disrupts Z-Disks and Leads to Myofibril Disassembly

GFP-labeled WT and NEECK α -actinin-2 were transiently expressed in NRCs for 18–48 hr.

(A) WT α -actinin shows normal Z-disk localization, and the titin T12 epitope is resolved as a single line in standard confocal microscopy.

(B) In contrast, NEECK leads to widening of the Z-disk and splitting of the T12 epitope after ~18 hr (asterisk); doublet T12 lines are highlighted by arrows.

(C) After 48 hr, Z-disks are completely disrupted and Z-disk titin, actin, and mutant α -actinin are localized in rod-like structures.

(D and E) Superresolution microscopy reveals that epitopes of N-terminal Z1Z2 of titin and their ligand telethonin are unresolvable in WT-transfected cells; NEECK causes widening of Z-disks. Doublet Z1Z2/telethonin lines are highlighted by arrows and the central α -actinin region is indicated by arrowheads. Insets show 2-fold enlargement. Z-disk titin (T12 epitope) or telethonin, red; mutant α -actinin-GFP, green; actin (Alexa 688-phalloidin) or titin Z1Z2, blue.

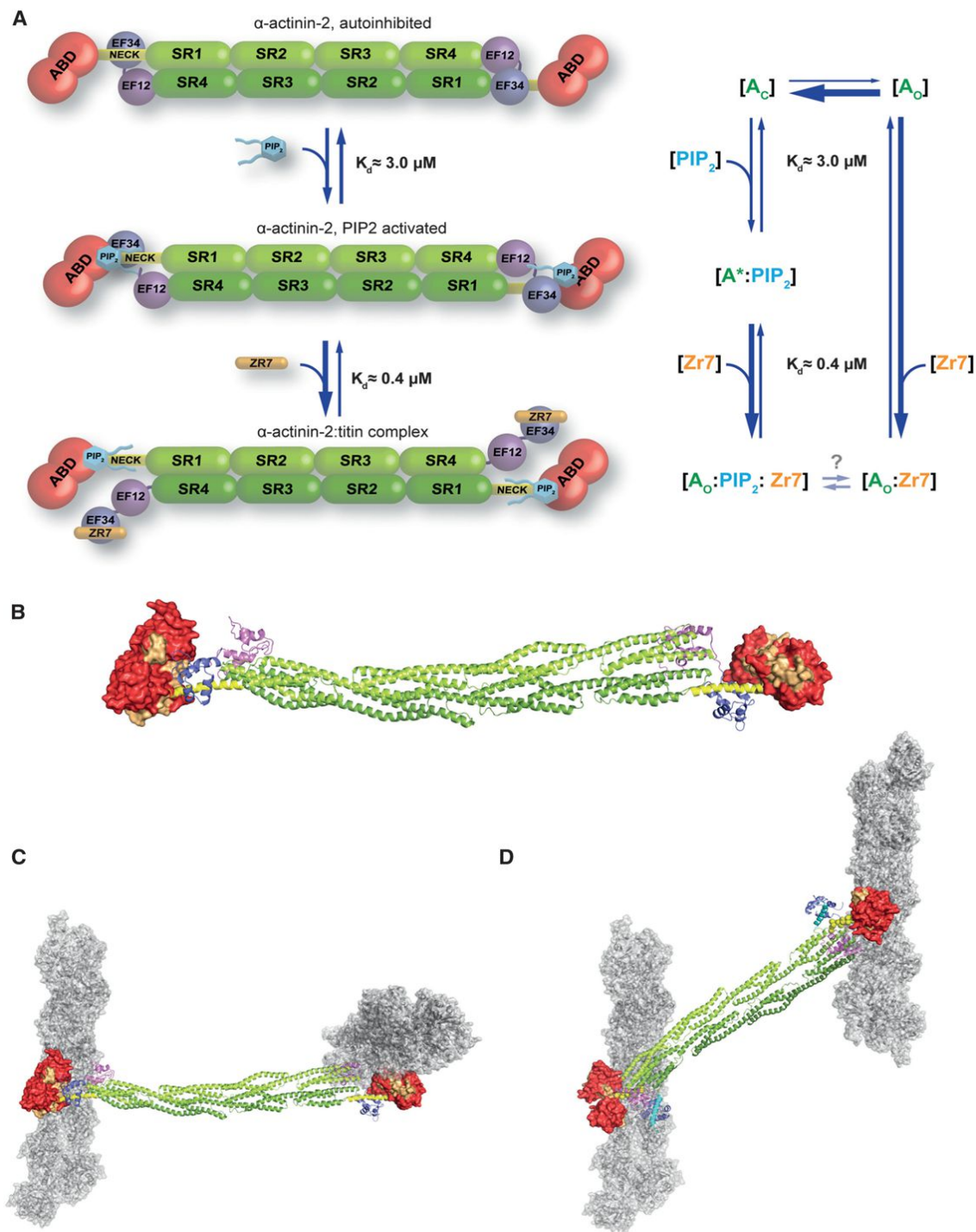


Figure 7. Molecular Mechanism of α -Actinin-2 Phosphoinositide-Based Activation and Model for F-Actin/ α -Actinin Crosslinking in the Z-Disk

(A) Reaction mechanism depicting α -actinin activation by PIP₂. α -actinin in the absence of PIP₂ is in equilibrium between highly populated closed [A_C] and low populated open states [A_O]. Addition of PIP₂ generates an activated state [$A^*:PIP_2$], with lower activation energy for opening. Binding of Zr-7 recruits α -actinin to the open conformation, leading to an increase of [$A_O:Zr-7$] due to higher α -actinin affinity for Zr-7 compared to the neck.

(B) Actin binding sites 1–3 (orange) mapped onto the molecular surface of ABD in α -actinin-2 in closed conformation. Color coding of domains is as in Figure 1.

(C) α -actinin-2 crystal structure superimposed on F-actin decorated by the CH1 domain of α -actinin-2 (PDB ID code 3LUE).

(D) Model of α -actinin-2 crosslinking antiparallel actin filaments. α -actinin-2 in open conformation (NEECK) was modeled assuming structural plasticity in the flexible neck, which allows for suitable orientation of ABDs. Titin Zr-7 is in cartoon presentation (cyan).

See also Movie S1.

CH1 domain (PDB ID code 3LUE) (Galkin et al., 2010). Notably, using the α -actinin-2 structure leads to a model with perpendicular actin filaments (Figure 7C), in strong disagreement with the established antiparallel actin architecture in the Z-disk but agreeing with the α -actinin-2 dimer architecture. Due to the internal twist of about 90° of the central rod (Ylänne et al., 2001), the ABDs in the dimer are rotated by 90°. We next used the structure of the NEECK mutant as the “open” structure in the Z-disk. Here the unbound neck is unstructured (Figure S4), allowing the ABD to explore different orientations, adopting those compatible with interaction with antiparallel actin filaments. Considering the angular distribution of F-actin and α -actinin, which centers at 60° and 120°, respectively (Hampton et al., 2007), we generated a model of two actin filaments crosslinked by an α -actinin-2 dimer (Figure 7D; Movie S1). In this model, the distance between filaments is ~230 Å, which is in excellent agreement with the observed interfilament distances in the tetragonal Z-disk lattice (240 Å; Goldstein et al., 1979).

The structural plasticity of α -actinin-2 has implications for not only its regulation but also for actin filament binding both in muscle and nonmuscle isoforms, where actin filaments are randomly oriented, requiring ABDs to adopt variable orientations. Interestingly, muscle α -actinin-2 was found to crosslink antiparallel as well as parallel actin filaments in α -actinin-F-actin rafts (Hampton et al., 2007). Assuming the absence of PIP2 in these assays, the flexibility of the ABD likely resides in the hinge region between the ABD and the neck (residue G258) (Figure 2A), similar to the solution structure of closed α -actinin-2 (Figure 4D). Other electron microscopy studies showed that ABDs in smooth muscle α -actinin can attain different orientations through movement in the flexible neck (Taylor and Taylor, 1993; Winkler et al., 1997), crosslinking both antiparallel and parallel actin filaments in vitro and in vivo (Meyer and Aebi, 1990; Tang et al., 2001).

Although alternative paths of titin in the Z-disk are conceivable, our results with NEECK show that even in strongly split Z-disks the Z1Z2 and telethonin epitopes remain in strict colocalization, implying that the two titin molecules crosslinked in the titin-telethonin complex must come from the same half-sarcomere, as suggested previously (Zou et al., 2006). Intriguingly, these findings also suggest that titin capping (via telethonin) and barbed-end actin filament capping by CapZ are not directly correlated, despite the close association of the titin Z1Z2-telethonin complex with the actin barbed end (Zou et al., 2006). However, actin capping by CapZ and crosslinking by α -actinin may crosstalk, as indeed α -actinin was reported to interact with CapZ via a binding site on the rod (Papa et al., 1999) and both proteins are PIP2 regulated (Figure S6A).

Impact of Pathogenic Mutations

Genetic variants in α -actinin genes are associated with several inherited diseases. Missense variants in nonmuscle actinin 1 (*ACTN1* gene) cause autosomal-dominant congenital macrothrombocytopenia (Guéguen et al., 2013; Kunishima et al., 2013), and approximately 4% of autosomal-dominant familial focal segmental glomerulosclerosis has been linked to nonmuscle *ACTN4* mutations (Kaplan et al., 2000). Missense variants in muscle *ACTN2* have been reported in sporadic cases and a few families with dilated or hypertrophic cardiomyopathy

(Chiu et al., 2010; Mohapatra et al., 2003; Theis et al., 2006). Our structure now provides a platform for the analysis of mutational impact on structure, ligand binding, and regulation of α -actinin in inherited human diseases (see Supplemental Information and Figure S7).

Structural Analysis of Selected Genetic Variants

The genetic variants are spread over all domains of α -actinin. Several variants are conservative and would not lead to major structural perturbations, in particular mutations in the CAMD and the rod domain. Changes of rod surface properties might, however, abrogate interactions with ligands, because the rod domain is recognized as the prominent protein interaction platform of α -actinin (Djinovic-Carugo et al., 2002). Interestingly, the mutations on the rod domain map on the less conserved, acidic side of the convex surface (Ylänne et al., 2001) (Figure S7). However, four mutations in the *ACTN1* and *ACTN4* genes have predicted disruptive potential: E225K (*ACTN1* gene, ABD) leads to a loss of a salt bridge and mutation R738W (*ACTN1* gene, CAMD) would disrupt the structure of the CAMD, whereas the W59R (*ACTN4* gene, ABD) and S262F (*ACTN4* gene, ABD) mutations destabilize the domain structure due to introduction of charged or bulky hydrophobic residues to the core of the ABD. Most hypertrophic cardiomyopathy and dilated cardiomyopathy variants were classified as structurally neutral.

Implication for Regulation of the α -Actinin Family in General

Muscle α -actinin interacts with many proteins via multiple binding sites. The CAMD EF3-4 site interacts with helical motifs in the actin- and α -actinin-binding proteins myopalladin, palladin, and myotilin, highly similar to the α -actinin-titin complex and the intramolecular neck complex detailed here (Beck et al., 2011). Dynamic regulation of α -actinin interactions with these proteins is therefore likely governed by the same principles as the one with titin. Additionally, the α -actinin-associated LIM protein (ALP) and ZASP/Cypher bind α -actinin at both the CAMD (via its PDZ domain) and the SR (Faulkner et al., 1999; Klaavuniemi et al., 2004). Although the binding sites for titin Zr-7 and the PDZ domain on CAMD do not coincide, an open structure might be required to accommodate both binding partners and prevent steric hindrance by the spatially close α -actinin domains.

Furthermore, interactions of CAMDs of the structurally related cytoskeletal actin-binding proteins dystrophin, utrophin, and spectrin may play important roles in regulating cytoskeletal interactions near the plasma membrane (Bennett and Healy, 2008), as suggested by recent studies on spectrin-ankyrin, actin, and protein 4.2 interactions (Korsgren and Lux, 2010; Korsgren et al., 2010). Although these EF hand domains retain aspects of calcium regulation (only the N-terminal EF hand binds calcium), the general mode of regulation seems highly similar to α -actinin, namely the CH2-R1 linker region of α/β -spectrin also binds to the CAMD EF3-4 hands, and this regulates protein interactions.

The mechanism we have detailed here is therefore likely to be of general relevance for regulating spectrin-like proteins via intramolecular pseudoligand interactions.

EXPERIMENTAL PROCEDURES

Purification and Crystallization

Proteins were expressed as His fusions in *Escherichia coli* and purified via Ni-NTA agarose and size-exclusion chromatography. Protein was lysine methylated and crystallized in a precipitant containing 0.2 M Mg formate, 5% PEG smear, and 10 mM EDTA by hanging-drop vapor diffusion at 14°C.

Structure Determination

A 3.5 Å data set was collected at beamline ID23-2 (European Synchrotron Radiation Facility [ESRF]). The phase problem was solved by molecular replacement using structures of the rod domain (PDB ID code 1HCI), the ABD from α -actinin-3 (PDB ID code 1WKU), and the NMR structure of EF3-4 (PDB ID code 1H8B) as search models.

Residues 34–892 were assigned in the final model. Details on data collection, processing, structure determination, and refinement are described in Extended Experimental Procedures and Table S1.

Electron Paramagnetic Resonance

Site-directed spin labeling (SDSL) was performed on native cysteine residues of WT and NEECK α -actinin-2. X band cw EPR experiments were carried out at 298 K or 160 K on a Bruker EMX spectrometer. Pulsed EPR measurements were carried out at 50 K on a Q band power upgraded Bruker ELEXSYS E580 spectrometer. Details are given in Extended Experimental Procedures.

SAXS Measurements and Modeling

Small-angle X-ray scattering data were collected at beamline X33 at European Molecular Biology Laboratory (EMBL) Hamburg for WT, NEECK, and PIP2 mutants at three different concentrations and analyzed following standard procedures. Molecular dynamics simulations were carried out using the GROMACS 4.0.7 package (Hess et al., 2008), whereas flexible docking was performed using GOLD version 5.2.2 (Jones et al., 1997). Further details are described in Extended Experimental Procedures.

Cell Biophysics

Experiments in neonatal rat cardiomyocytes were performed using published methods and antibodies using live-cell imaging on a Zeiss LSM510 confocal microscope and superresolution on a Leica TCS STED instrument (see Extended Experimental Procedures).

ACCESSION NUMBERS

The coordinates and structure factors of the α -actinin-2 structure have been deposited in the Protein Data Bank under ID code 4D1E.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, five tables, and one movie and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2014.10.056>.

AUTHOR CONTRIBUTIONS

Experiments were designed by M.G., K.D.-C., and E.d.A.R. in consultation with K.F.P., E.B., and D.I.S.; all structural work was performed by E.d.A.R., N.P., A.S., B.S., K.F.P., and P.V.K.; E.d.A.R., A.G., and J.K. performed biochemical work; A.G. and M.R.H. performed cell biophysics, and F.L.A. performed the NMR work; C.S. and E.A.G. assisted with protein purification and crystallization; A.A.P. and B.Ž. performed MD; and M.G., K.D.-C., E.d.A.R., and N.P. wrote the manuscript, with all authors contributing to editing the manuscript and supporting the conclusions.

ACKNOWLEDGMENTS

We are greatly indebted to Ay Lin Kho for NRC preparations. Thanks also to Dusan Turk (Institute Jozef Stefan, Ljubljana) for initial help with refinement

and Oliviero Carugo (University of Vienna and University of Pavia) and Bettina Hartlieb (Baxter Innovations GmbH) for critical reading of the manuscript. We thank the staff of the MX beamlines at the ESRF in Grenoble, SAXS beamline X33 (Deutsches Elektronen-Synchrotron, Hamburg), and SWING beamline (Soleil, Saint-Aubin) for their excellent support. We thank the staff of Campus Science Support Facilities GmbH (Campus Vienna Biocenter) for technical assistance. The K.D.-C. group was supported by Austrian Science Fund (FWF) Projects I525, I1593, P22276, and P19060, by the Federal Ministry of Economy, Family and Youth through the initiative “Laura Bassi Centres of Expertise” funding the Center of Optimized Structural Studies (253275), by the Marie Curie Initial Training Network: MUZIC (238423), and by the University of Vienna. This research was also funded by the European Community Seventh Framework Programme (FP7/2007–2013) under BioStruct-X (283570). F.L.A. thanks BioTek2021 Project 217708/O10 and the Research Council of Norway for financial support. B.Ž. and A.A.P. were supported by the European Research Council (279408). A.A.P. was supported by the Russian Scientific Foundation (14-24-00118). M.G. and A.G. were supported by the Leducq Foundation Transatlantic Network of Excellence: Proteotoxicity (11 CVD 04) and the Medical Research Council of Great Britain (MR/J010456/1). A.G. and E.d.A.R. were also supported by the Marie Curie Initial Training Network: MUZIC (238423). M.G. holds the British Heart Foundation Chair of Molecular Cardiology (CH/08/001). D.I.S. acknowledges support from the Human Frontier Science Program (RGP0017/2012).

Received: May 7, 2014

Revised: October 1, 2014

Accepted: October 24, 2014

Published: November 26, 2014

REFERENCES

- Atkinson, R.A., Joseph, C., Kelly, G., Muskett, F.W., Frenkiel, T.A., Nietlispach, D., and Pastore, A. (2001). Ca^{2+} -independent binding of an EF-hand domain to a novel motif in the α -actinin-titin complex. *Nat. Struct. Biol.* 8, 853–857.
- Bayley, P.M., Findlay, W.A., and Martin, S.R. (1996). Target recognition by calmodulin: dissecting the kinetics and affinity of interaction using short peptide sequences. *Protein Sci.* 5, 1215–1228.
- Beck, M.R., Otey, C.A., and Campbell, S.L. (2011). Structural characterization of the interactions between palladin and α -actinin. *J. Mol. Biol.* 413, 712–725.
- Bennett, V., and Healy, J. (2008). Organizing the fluid membrane bilayer: diseases linked to spectrin and ankyrin. *Trends Mol. Med.* 14, 28–36.
- Bernadó, P., Mylonas, E., Petoukhov, M.V., Blackledge, M., and Svergun, D.I. (2007). Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.* 129, 5656–5664.
- Carugo, O. (2014). Wolumes - an algorithm to compute the volume of atoms and residues in proteins. *arXiv*, <http://arxiv.org/pdf/1406.3242.pdf>.
- Chin, D., and Means, A.R. (2000). Calmodulin: a prototypical calcium sensor. *Trends Cell Biol.* 10, 322–328.
- Chiu, C., Bagnall, R.D., Ingles, J., Yeates, L., Kennerson, M., Donald, J.A., Jormakka, M., Lind, J.M., and Semsarian, C. (2010). Mutations in α -actinin-2 cause hypertrophic cardiomyopathy: a genome-wide analysis. *J. Am. Coll. Cardiol.* 55, 1127–1135.
- Djinovic-Carugo, K., Gautel, M., Ylännä, J., and Young, P. (2002). The spectrin repeat: a structural platform for cytoskeletal protein assemblies. *FEBS Lett.* 513, 119–123.
- Faulkner, G., Pallavicini, A., Formentin, E., Comelli, A., Ievolelella, C., Trevisan, S., Bortoletto, G., Scannapieco, P., Salamon, M., Mouly, V., et al. (1999). ZASP: a new Z-band alternatively spliced PDZ-motif protein. *J. Cell Biol.* 146, 465–475.
- Foley, K.S., and Young, P.W. (2014). The non-muscle functions of actinins: an update. *Biochem. J.* 459, 1–13.
- Fraleigh, T.S., Tran, T.C., Corgan, A.M., Nash, C.A., Hao, J., Critchley, D.R., and Greenwood, J.A. (2003). Phosphoinositide binding inhibits α -actinin bundling activity. *J. Biol. Chem.* 278, 24039–24045.

- Franzot, G., Sjöblom, B., Gautel, M., and Djinić Carugo, K. (2005). The crystal structure of the actin binding domain from α -actinin in its closed conformation: structural insight into phospholipid regulation of α -actinin. *J. Mol. Biol.* 348, 151–165.
- Fukami, K., Furuhashi, K., Inagaki, M., Endo, T., Hatano, S., and Takenawa, T. (1992). Requirement of phosphatidylinositol 4,5-bisphosphate for α -actinin function. *Nature* 359, 150–152.
- Fukami, K., Sawada, N., Endo, T., and Takenawa, T. (1996). Identification of a phosphatidylinositol 4,5-bisphosphate-binding site in chicken skeletal muscle α -actinin. *J. Biol. Chem.* 271, 2646–2650.
- Full, S.J., Deinzer, M.L., Ho, P.S., and Greenwood, J.A. (2007). Phosphoinositide binding regulates α -actinin CH2 domain structure: analysis by hydrogen/deuterium exchange mass spectrometry. *Protein Sci.* 16, 2597–2604.
- Galkin, V.E., Orlova, A., Salmazo, A., Djinić-Carugo, K., and Egelman, E.H. (2010). Opening of tandem calponin homology domains regulates their affinity for F-actin. *Nat. Struct. Mol. Biol.* 17, 614–616.
- Gautel, M. (2011). The sarcomeric cytoskeleton: who picks up the strain? *Curr. Opin. Cell Biol.* 23, 39–46.
- Gautel, M., Goulding, D., Bullard, B., Weber, K., and Fürst, D.O. (1996). The central Z-disk region of titin is assembled from a novel repeat in variable copy numbers. *J. Cell Sci.* 109, 2747–2754.
- Goldstein, M.A., Schroeter, J.P., and Sass, R.L. (1979). The Z lattice in canine cardiac muscle. *J. Cell Biol.* 83, 187–204.
- Goldstein, M.A., Michael, L.H., Schroeter, J.P., and Sass, R.L. (1988). Structural states in the Z band of skeletal muscle correlate with states of active and passive tension. *J. Gen. Physiol.* 92, 113–119.
- Guéguen, P., Rouault, K., Chen, J.M., Raguénès, O., Fichou, Y., Hardy, E., Gobin, E., Pan-Petes, B., Kerbirou, M., Trouvé, P., et al. (2013). A missense mutation in the α -actinin 1 gene (ACTN1) is the cause of autosomal dominant macrothrombocytopenia in a large French family. *PLoS ONE* 8, e74728.
- Hampton, C.M., Taylor, D.W., and Taylor, K.A. (2007). Novel structures for α -actinin:F-actin interactions and their implications for actin-membrane attachment and tension sensing in the cytoskeleton. *J. Mol. Biol.* 368, 92–104.
- Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008). GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* 4, 435–447.
- Hoeftlich, K.P., and Ikura, M. (2002). Calmodulin in action: diversity in target recognition and activation mechanisms. *Cell* 108, 739–742.
- Jones, G., Willett, P., Glen, R.C., Leach, A.R., and Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* 267, 727–748.
- Kaplan, J.M., Kim, S.H., North, K.N., Rennke, H., Correia, L.A., Tong, H.Q., Mathis, B.J., Rodríguez-Pérez, J.C., Allen, P.G., Beggs, A.H., and Pollak, M.R. (2000). Mutations in ACTN4, encoding α -actinin-4, cause familial focal segmental glomerulosclerosis. *Nat. Genet.* 24, 251–256.
- Klaavuniemi, T., Kelloniemi, A., and Ylänné, J. (2004). The ZASP-like motif in actinin-associated LIM protein is required for interaction with the α -actinin rod and for targeting to the muscle Z-line. *J. Biol. Chem.* 279, 26402–26410.
- Klein, M.G., Shi, W., Ramagopal, U., Tseng, Y., Wirtz, D., Kovar, D.R., Staiger, C.J., and Almo, S.C. (2004). Structure of the actin crosslinking core of fimbrin. *Structure* 12, 999–1013.
- Korsgren, C., and Lux, S.E. (2010). The carboxyterminal EF domain of erythroid α -spectrin is necessary for optimal spectrin-actin binding. *Blood* 116, 2600–2607.
- Korsgren, C., Peters, L.L., and Lux, S.E. (2010). Protein 4.2 binds to the carboxyl-terminal EF-hands of erythroid α -spectrin in a calcium- and calmodulin-dependent manner. *J. Biol. Chem.* 285, 4757–4770.
- Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774–797.
- Kunishima, S., Okuno, Y., Yoshida, K., Shiraishi, Y., Sanada, M., Muramatsu, H., Chiba, K., Tanaka, H., Miyazaki, K., Sakai, M., et al. (2013). ACTN1 mutations cause congenital macrothrombocytopenia. *Am. J. Hum. Genet.* 92, 431–438.
- Li, J., and Russell, B. (2013). Phosphatidylinositol 4,5-bisphosphate regulates CapZ β 1 and actin dynamics in response to mechanical strain. *Am. J. Physiol. Heart Circ. Physiol.* 305, H1614–H1623.
- Luther, P.K. (2009). The vertebrate muscle Z-disc: sarcomere anchor for structure and signalling. *J. Muscle Res. Cell Motil.* 30, 171–185.
- Masaki, T., Endo, M., and Ebashi, S. (1967). Localization of 6S component of α -actinin at Z-band. *J. Biochem.* 62, 630–632.
- Meyer, R.K., and Aebi, U. (1990). Bundling of actin filaments by α -actinin depends on its molecular length. *J. Cell Biol.* 110, 2013–2024.
- Mohapatra, B., Jimenez, S., Lin, J.H., Bowles, K.R., Coveler, K.J., Marx, J.G., Chrisco, M.A., Murphy, R.T., Lurie, P.R., Schwartz, R.J., et al. (2003). Mutations in the muscle LIM protein and α -actinin-2 genes in dilated cardiomyopathy and endocardial fibroelastosis. *Mol. Genet. Metab.* 80, 207–215.
- Papa, I., Astier, C., Kwiatek, O., Raynaud, F., Bonnal, C., Lebart, M.C., Roustan, C., and Benyamin, Y. (1999). α -actinin-CapZ, an anchoring complex for thin filaments in Z-line. *J. Muscle Res. Cell Motil.* 20, 187–197.
- Saad, J.S., Miller, J., Tai, J., Kim, A., Ghanam, R.H., and Summers, M.F. (2006). Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *Proc. Natl. Acad. Sci. USA* 103, 11364–11369.
- Sanger, J.M., and Sanger, J.W. (2008). The dynamic Z bands of striated muscle cells. *Sci. Signal.* 1, pe37.
- Schacht, J. (1976). Inhibition by neomycin of polyphosphoinositide turnover in subcellular fractions of guinea-pig cerebral cortex in vitro. *J. Neurochem.* 27, 1119–1124.
- Sjöblom, B., Salmazo, A., and Djinić-Carugo, K. (2008). α -actinin structure and regulation. *Cell. Mol. Life Sci.* 65, 2688–2701.
- Sorimachi, H., Freiburg, A., Kolmerer, B., Ishiura, S., Stier, G., Gregorio, C.C., Labeit, D., Linke, W.A., Suzuki, K., and Labeit, S. (1997). Tissue-specific expression and α -actinin binding properties of the Z-disc titin: implications for the nature of vertebrate Z-discs. *J. Mol. Biol.* 270, 688–695.
- Swindells, M.B., and Ikura, M. (1996). Pre-formation of the semi-open conformation by the apo-calmodulin C-terminal domain and implications for binding IQ-motifs. *Nat. Struct. Biol.* 3, 501–504.
- Takeda, S., Yamashita, A., Maeda, K., and Maeda, Y. (2003). Structure of the core domain of human cardiac troponin in the Ca(2+)-saturated form. *Nature* 424, 35–41.
- Tang, J., Taylor, D.W., and Taylor, K.A. (2001). The three-dimensional structure of α -actinin obtained by cryoelectron microscopy suggests a model for Ca(2+)-dependent actin binding. *J. Mol. Biol.* 310, 845–858.
- Taylor, K.A., and Taylor, D.W. (1993). Projection image of smooth muscle α -actinin from two-dimensional crystals formed on positively charged lipid layers. *J. Mol. Biol.* 230, 196–205.
- Theis, J.L., Bos, J.M., Bartleson, V.B., Will, M.L., Binder, J., Vatta, M., Towbin, J.A., Gersh, B.J., Ommen, S.R., and Ackerman, M.J. (2006). Echocardiographic-determined septal morphology in Z-disc hypertrophic cardiomyopathy. *Biochem. Biophys. Res. Commun.* 351, 896–902.
- Tskhovrebova, L., and Trinick, J. (2010). Roles of titin in the structure and elasticity of the sarcomere. *J. Biomed. Biotechnol.* 2010, 612482.
- Winkler, J., Lünsdorf, H., and Jockusch, B.M. (1997). Flexibility and fine structure of smooth-muscle α -actinin. *Eur. J. Biochem.* 248, 193–199.
- Ylänné, J., Scheffzek, K., Young, P., and Saraste, M. (2001). Crystal structure of the α -actinin rod reveals an extensive torsional twist. *Structure* 9, 597–604.
- Young, P., and Gautel, M. (2000). The interaction of titin and α -actinin is controlled by a phospholipid-regulated intramolecular pseudoligand mechanism. *EMBO J.* 19, 6331–6340.
- Young, P., Ferguson, C., Bañuelos, S., and Gautel, M. (1998). Molecular structure of the sarcomeric Z-disc: two types of titin interactions lead to an asymmetrical sorting of α -actinin. *EMBO J.* 17, 1614–1624.
- Zou, P., Pinotsis, N., Lange, S., Song, Y.H., Popov, A., Mavridis, I., Mayans, O.M., Gautel, M., and Wilmanns, M. (2006). Palindromic assembly of the giant muscle protein titin in the sarcomeric Z-disc. *Nature* 439, 229–233.

Integration of Genomic Data Enables Selective Discovery of Breast Cancer Drivers

Félix Sanchez-Garcia,^{1,2,5} Patricia Villagrasa,^{3,5} Junji Matsui,³ Dylan Kotliar,¹ Verónica Castro,³ Uri-David Akavia,^{1,4} Bo-Juen Chen,¹ Laura Saucedo-Cuevas,³ Ruth Rodriguez Barrueco,³ David Llobet-Navas,³ Jose M. Silva,^{3,6,*} and Dana Pe'er^{1,6,*}

¹Department of Biological Sciences and Department of Systems Biology, Columbia University, New York, NY 10027, USA

²Department of Computer Science, Columbia University, New York, NY 10027, USA

³Icahn School of Medicine at Mount Sinai, The Mount Sinai Hospital, New York, NY 10029, USA

⁴Present address: Department of Biochemistry, Faculty of Medicine, McGill University, Montréal, QC H3A 0G4, Canada

⁵Co-first author

⁶Co-senior author

*Correspondence: jose.silva@mssm.edu (J.M.S.), dpeer@biology.columbia.edu (D.P.)

<http://dx.doi.org/10.1016/j.cell.2014.10.048>

SUMMARY

Identifying driver genes in cancer remains a crucial bottleneck in therapeutic development and basic understanding of the disease. We developed Helios, an algorithm that integrates genomic data from primary tumors with data from functional RNAi screens to pinpoint driver genes within large recurrently amplified regions of DNA. Applying Helios to breast cancer data identified a set of candidate drivers highly enriched with known drivers ($p < 10^{-14}$). Nine of ten top-scoring Helios genes are known drivers of breast cancer, and in vitro validation of 12 candidates predicted by Helios found ten conferred enhanced anchorage-independent growth, demonstrating Helios's exquisite sensitivity and specificity. We extensively characterized RSF-1, a driver identified by Helios whose amplification correlates with poor prognosis, and found increased tumorigenesis and metastasis in mouse models. We have demonstrated a powerful approach for identifying driver genes and how it can yield important insights into cancer.

INTRODUCTION

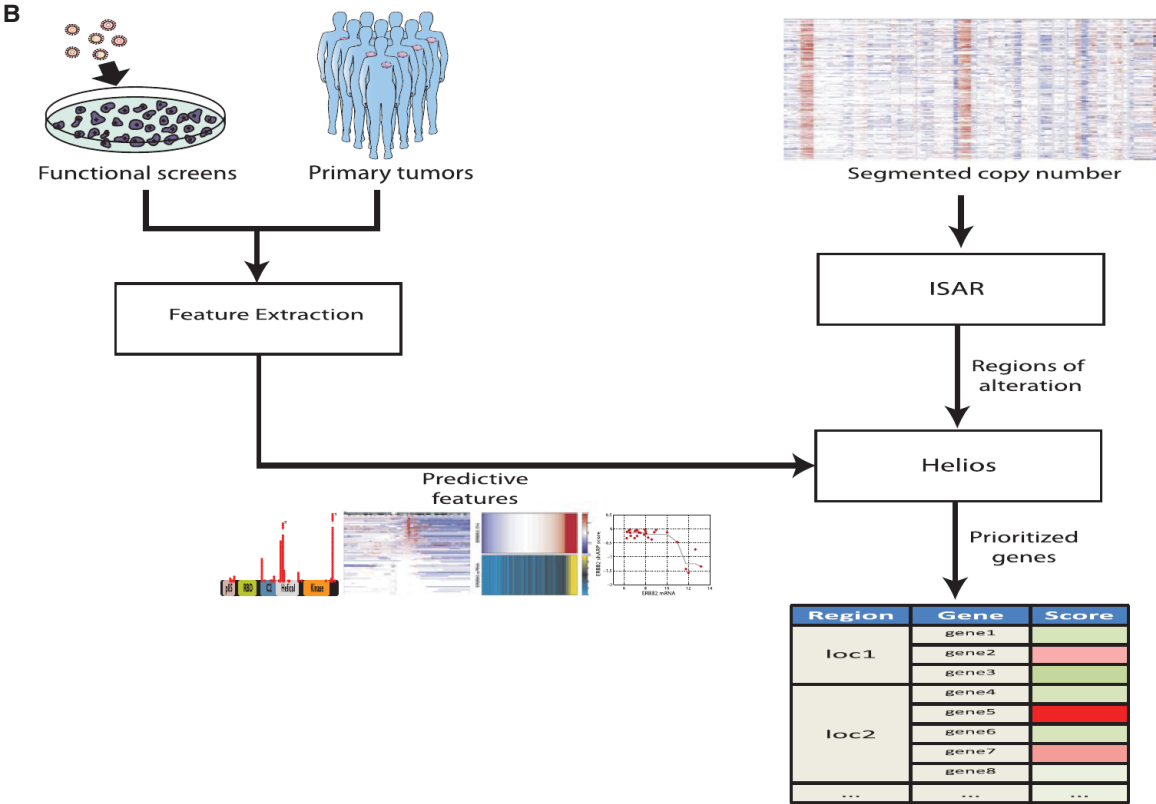
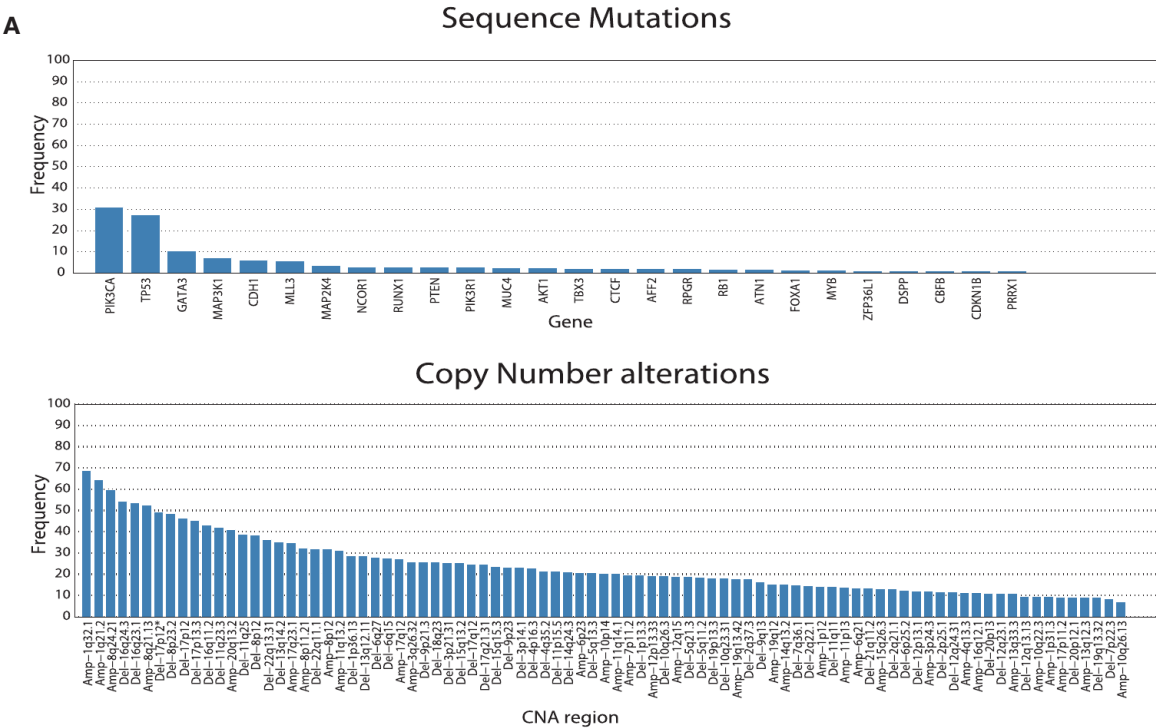
Cancer genome data collected by projects such as the The Cancer Genome Atlas (TCGA) or the International Cancer Genome Consortium (ICGC) is defining the landscape of genetic alterations that underlie cancer. Tumor cells may harbor thousands of genetic lesions including point mutations, somatic copy-number alterations, and translocations that localize to hundreds or even thousands of genes. However, most affected genes are so-called passengers and their alteration does not confer any type of advantage to tumors (Vogelstein et al., 2013). A pivotal challenge in cancer genomics is to identify the small subset of altered genes (so-called drivers) that directly contribute to tumor fitness and progression.

Exome sequencing studies helped identify driver genes (Curtis et al., 2012; Stephens et al., 2012), however, the majority of point

mutations display low population frequencies, with only a handful altered in >5% of patients (Stephens et al., 2012). In breast cancer, only six genes have point mutations in >5% of samples, and of these, only PIK3CA (36% frequency) is currently targeted therapeutically (Cancer Genome Atlas Network, 2012). Instead, the most recurrent genetic lesions in breast cancer are somatic copy-number alterations (SCNAs), often driven by inactivation of DNA repair genes such as BRCA1/2. Indeed, HER2, one of the most therapeutically targeted drivers in breast cancer, is primarily dysregulated by copy-number amplification.

The ability to discern drivers from copy-number alteration promises to dramatically expand the set of therapeutic targets in this disease. However, this potential is crucially hindered by the difficulty of driver discovery (Yuan et al., 2012). The crux of the difficulty is that in all but a few instances, these lesions contain dozens of genes and no previously characterized drivers (Albertson et al., 2003). A recent study analyzing multiple tumor types reported that over 70% of 140 recurrently altered regions did not contain a known oncogene or tumor suppressor (Zack et al., 2013). As a result, most recent driver discovery efforts have focused on point mutations, which directly indicate the target genes by virtue of their precise location (Kandoth et al., 2013; Lohr et al., 2012; Wong et al., 2011), and less progress has been made with respect to SCNAs. However, the increased frequency of recurring SCNAs relative to point mutations (87 SCNA regions versus six mutated genes with >5% population frequency) (Figure 1A) highlights the need for methods to pinpoint drivers within these regions.

Genome-wide pooled-RNAi screening is an alternative approach to driver gene discovery. In these studies, a short hairpin RNA (shRNA) library is transduced into cancer cell lines and the growth effect of each individual gene knockdown is assessed for each cell line (Cheung et al., 2011; Marcotte et al., 2012; Silva et al., 2008). While such studies can provide gene-level resolution, they are currently limited by the high degree of noise, the potential for off-target effects of shRNAs and by the artificiality of the in vitro screening system (Kaelin, 2012). Moreover, cell-lines are not fully representative of primary tumor biology as these lack tissue structure and microenvironment, which are key to cellular behavior (Bissell and Hines, 2011).



(legend on next page)

Given the largely orthogonal strengths and weaknesses of descriptive analysis of primary cancer genomes and in vitro genome-wide functional screening, we hypothesized that integrating the two data types into a single approach would result in increased resolution and accuracy for driver gene discovery. Therefore, we developed Helios (Figure 1B), an algorithm that incorporates primary tumor SCNA, point mutations, gene expression, and RNAi screens into a single candidate driver score. Helios runs in two steps, first identifying regions of focal SCNAs and then identifying driver genes within each region by integrating functional screens and other data using a Bayesian transfer-learning framework.

Helios displayed a remarkable capacity to pinpoint bona fide cancer drivers when the algorithm was used to analyze the SCNA landscape of breast cancer. In a systematic evaluation of Helios's performance, we selected 12 driver candidates identified by Helios, based on their frequency of occurrence, for experimental investigation. We found ten of 12 candidate genes induced increased anchorage-independent growth when overexpressed in vitro. Thus, Helios demonstrated an unprecedented sensitivity and specificity in identifying genes that promote oncogenic capabilities. Helios doubled the number of SCNA drivers identified in breast cancer and substantially increased our understanding of the breast cancer SCNA landscape.

RESULTS

ISAR Expands the List of Significantly Amplified Regions in Breast Cancer

The first step for identifying SCNA-drivers is identification of significantly altered regions. There are multiple algorithms that successfully perform this task (Mermel et al., 2011; Walter et al., 2011), GISTIC2 being the most widely used among these. We noted a number of oncogenes (e.g., *BCL2*) that were not detected as falling within a significantly altered region by GISTIC2 in the TCGA breast cancer data (Cancer Genome Atlas Network, 2012). By visual inspection of chromosome 18, we noted that while *BCL2* does not appear significantly amplified based on its absolute copy-level, its copy number is nevertheless significantly higher than the adjacent chromosomal regions (Figure S1 available online). Most SCNA detection algorithms, including GISTIC2, compute a null distribution across the entire genome to estimate the significance of alterations. However, the alteration rate can strongly differ across different genomic regions, due to features such as DNA secondary structure and DNA hypomethylation (De and Michor, 2011).

Therefore, we developed Identification of Significantly Altered Regions (ISAR), an algorithm that accounts for local differences in SCNA rate due to these and other forces. By computing the

significance locally, the algorithm is capable of identifying both global alteration events, as well as subtle events, such as a focal amplification within largely deleted regions, that would be missed if the background distribution for the whole genome were employed (see Experimental Procedures). We applied ISAR to 785 breast cancer samples (Cancer Genome Atlas Network, 2012) and identified 83 significantly amplified regions (see Table S1), compared to the 30 regions originally reported by the TCGA consortium. ISAR captures all significant regions captured by GISTIC2 and many additional regions. Among the new regions, we find many bona fide or likely oncogenes, including *MYB*, *BCL2*, *CDK4*, *ESR1*, *FGFR2*, *FGFR3*, and *FGFR4*. Identified regions contained an average of 14 genes resulting in a total of 1,226 significantly amplified genes across all 83 regions.

Helios: An Integrative Approach to Pinpoint Drivers

Helios seeks to exploit additional properties—e.g., recurrent domain-specific point mutations or depletion in a lethality shRNA screen—to implicate likely driver genes targeted by the SCNA. Helios considers the entire significantly altered region, but prioritizes the genes within this region by incorporating cues from additional genetic and genomic data to estimate the probability that each gene is a driver (Figure 2A). It is a statistically rigorous framework for combining multiple signals that might lack power individually into a single score for the likelihood that each gene's amplification specifically increases tumor fitness. Here, we integrate features derived from exome-sequencing, shRNA screening, and gene-expression, but due to the flexibility of our framework, these could readily be removed, modified, or extended for subsequent studies.

Helios uses a set of features to classify genes as either drivers or passengers, based on inference within a hierarchical Bayesian mixture model (see Experimental Procedures; Figure S2). Standard classification approaches rely on an initial list of examples—drivers and passengers—to train the model. Unfortunately, the list of known oncogenic drivers is relatively small and strongly biased toward kinases and extreme phenotypes that facilitate discovery. Instead, Helios begins with the assumption that a driver gene is more likely to be near the most frequently amplified segment (defined as peak) of the ISAR region. This is used to initialize the algorithm by providing an estimated list of drivers to start from. Helios then iterates between two stages until convergence by (1) learning the parameters to distinguish passengers and drivers on the basis of their SCNA profile and on the additional genomic data, and (2) recomputing the probability that each gene is a driver using the parameters determined in step 1.

Helios uses a transfer learning approach (Widmer and Rätsch, 2011) whereby drivers with clearer signal (e.g., at the peak of

Figure 1. Helios Integrates Data from Primary Tumor and Functional Screens

(A) Frequency of alteration in the TCGA breast cancer data of (top) genes with recurrent point mutations and (bottom) regions of recurrent copy-number alteration. Significant genes and regions were downloaded from the DBroad Genome Data analysis Center, selecting the TCGA pipeline algorithms GISTIC2 (v. 4.2012021700.0.0) and MutSig (v. 4.2011112800.0.0).

(B) A schematic of our pipeline for the identification of candidate driver genes. The method first uses ISAR to identify regions of focal SCNAs. To pinpoint drivers within those regions, it extracts features from genetic, genomic and functional data, which are integrated into a single probabilistic score by Helios.

See also Figure S1 and Table S1.

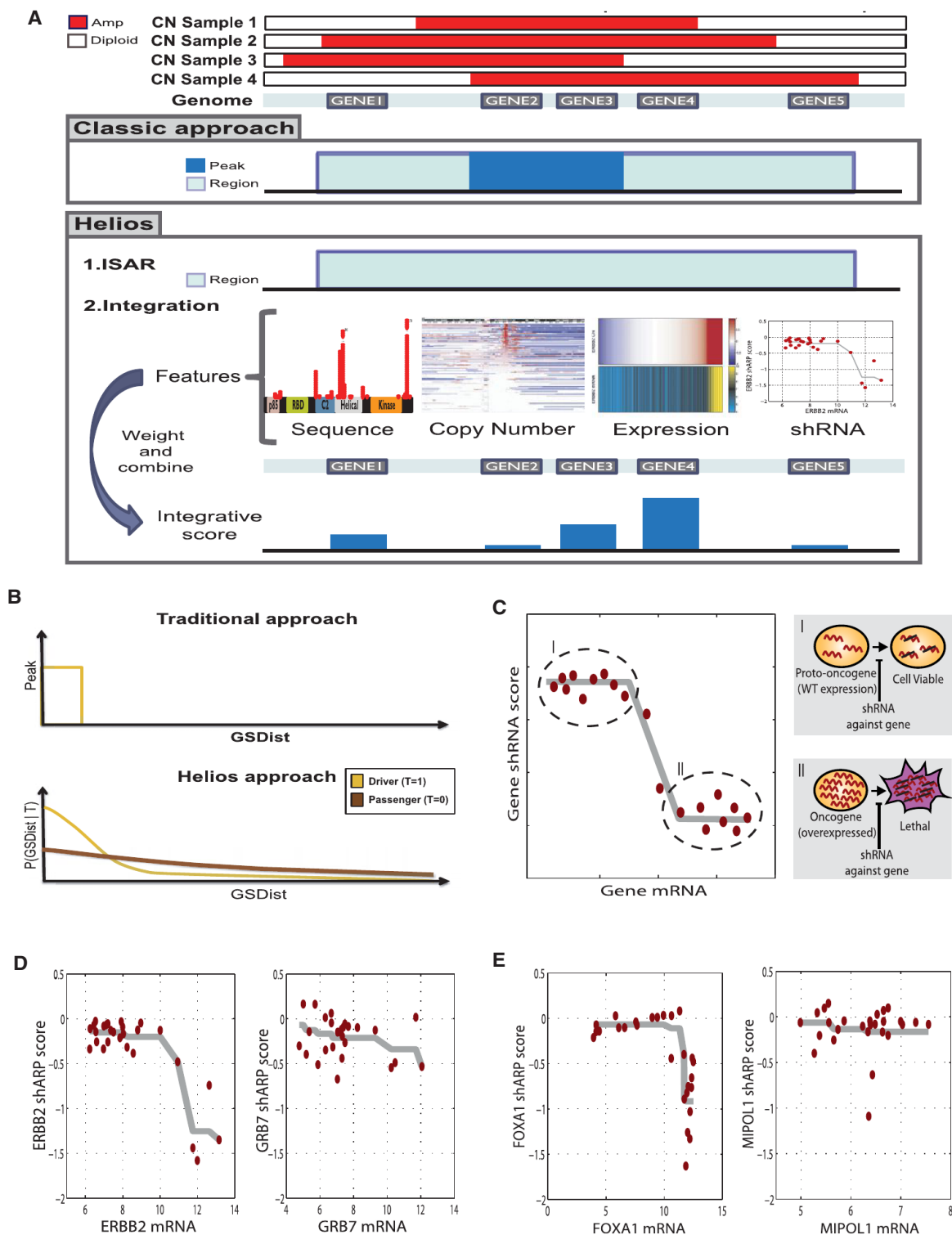


Figure 2. Helios Features

(A) Diagram of the classic and Helios approach. While the classic approach relies solely on copy number, both to identify significantly altered regions and to further narrow down those region to a minimal region of maximal alteration, Helios identifies regions in the same fashion, but then integrates features extracted from different data sources to compute the probability of each gene being a target of the region.

(legend continued on next page)

their region) are used to extract informative features to improve performance in cases with less obvious signal. Helios automatically learns the weights of features directly from the data by leveraging information among features. In each iteration, Helios learns a better classification of drivers and passengers, which in turn is used to learn better parameters, until convergence (see Extended Experimental Procedures). Helios utilizes a mixture of two copy-number distributions—one for drivers and one for passengers, thus avoiding the problematic selection of a hard threshold for defining aberrant regions (Figure 2B). Additionally, Helios permits final models where more than one gene in a region is identified as a likely driver, or where no probable driver genes are identified.

Finally, Helios can readily incorporate additional features, including complex features generated by combinations of multiple data sources. It automatically learns the contribution and importance of each feature directly from the data, making it easily extendable and adaptable to other cancer types. For example, here, we integrate data from functional screens based on the concept of oncogene addiction (Weinstein and Joe, 2008) by deriving a composite statistic reflecting the extent to which shRNA-depletion in a genome-wide screen correlated with overexpression of the gene at baseline. A similar idea has recently been used to discover the oncogene *HNF1B* (Shao et al., 2013). Our oncogene addiction score allows for both linear and nonlinear relations between gene expression and lethality (See Figures 2B–2D; Experimental Procedures). This ability to combine multiple weaker pieces of evidence from heterogeneous data types into a single score enables Helios to effectively pinpoint the driver gene from within the recurrently altered region.

Helios Identifies Candidate Drivers of Breast Cancer

We used Helios to integrate TCGA data from 785 primary breast cancer tumors, including DNA copy number, gene expression, and sequence mutations (Cancer Genome Atlas Network, 2012), with data from 27 breast cancer cell lines including gene expression, copy number, and shRNA depletion in a genome-wide shRNA screen (Barretina et al., 2012; Marcotte et al., 2012).

Using stringent criteria, we defined 64 candidate drivers by selecting only the top gene in each region and applying a threshold of Helios score >0.5 (see Table S2). Some significant SCNA regions did not contain a high scoring protein-coding gene; these amplifications potentially target noncoding RNA or other genomic features. For example, all protein-coding genes were low scoring in an amplified region containing the known oncomir *mir21* (O'Day and Lal, 2010). While ~20% of the regions con-

tained more than one high scoring gene, we limited our initial analysis to the highest scoring gene in each region.

To evaluate the sensitivity of our approach, we combined several publically available resources to create a comprehensive set of breast cancer oncogenes (Beroukheim et al., 2010; Uniprot Consortium, 2013; Pletscher-Frankild et al., 2014) (Figure S3; Extended Experimental Procedures). Among the ten top-scoring Helios genes, nine were included in this set (*FOXA1*, *PIK3CA*, *CCND1*, *CDK4*, *MYB*, *ERBB2*, *IGF1R*, *BCL2*, and *ESR1*), while only five of these appear in regions that are significant based on GISTIC2. Moreover, the entire list of 64 Helios candidates was significantly enriched for our compiled set of breast cancer drivers (16/64, p value < 4×10^{-15}), a large improvement over the set of all genes in amplified regions identified by GISTIC2 (17/452, p value > 10^{-3}) (Cancer Genome Atlas Network, 2012) (Figure 3A). The performance of the method was also compared against two other algorithms, GAIA (Morganella et al., 2011) and DiNAMIC (Walter et al., 2011), outperforming both of them (18/768, p value > 10^{-3} and 185/10,651, p value > 10^{-3} , respectively). This demonstrates the significant improvement of our integrative approach over the state of the art.

Helios's integration across multiple data sources is key to its ability to be both specific and sensitive. Sequence mutations are gene-specific, but only few drivers harbor such mutations recurrently. SCNAs typically cover a large number of genes, making it hard to identify the target of the amplification based on copy number alone. For instance, *CDK4* shares exactly the same copy-number profile with its five closest neighbors, but the lethality displayed by *CDK4* in the shRNA screen raises its Helios score (Figure 3B). More strikingly, *BCL2* is only the sixth gene in its region in terms of copy-number alteration frequency, but its dramatic oncogene addiction score raises its Helios score well above all others in the region (Figure 3C). In many cases (e.g., *EGFR* or *ADAM15*) (Figures 3D and 3E), it is not any single feature, but a combination of features that identifies the top-scoring gene in the region. Figure 3F shows how Helios outperforms the simple use of the data sources independently to identify drivers. Even if all of the candidates obtained by each data source are joined together naively, Helios provides significantly better sensitivity (15 versus 9 detected driver genes) and specificity (hypergeometric enrichment p value of driver genes 8.16×10^{-14} versus 4.72×10^{-11}).

Candidate Selection for Systematic In Vitro Validation of Helios-Predicted Genes

Helios is designed to rank genes within an amplified region based on their likely driver capacity. Contrary to most prior

(B) Diagram of the copy-number model of the Helios Algorithm. The classic approach (top) calculates a hard threshold on the delta to the most altered marker (GSDist, x axis) to define the peak region (y axis). Helios (bottom) instead calculates the probability (y axis) of displaying a GSDist value (x axis) for both driver and passenger genes (yellow or brown curves respectively).

(C) Our oncogene addiction score uses monotonic regression to measure the association between gene dosage (x axis) and shRNA dropout (y axis), aiming to differentiate the proto-oncogenic state (I) of the driver, which is expressed at wild-type levels, and the oncogenic state (II), which is characterized by high expression and high dependency on the gene for survival.

(D) Monotonic regression of the shRNA dropout (y axis) based on the gene dosage (x axis) for the top-scoring gene for oncogene addiction in the 17q12 region (left) and its neighboring gene in the genome (right). (E) Monotonic regression of the shRNA dropout (y axis) based on the gene dosage (x axis) for the top-scoring gene for oncogene addiction in the 14q13 region (left) and its neighboring gene in the genome (right).

See also Figure S2.

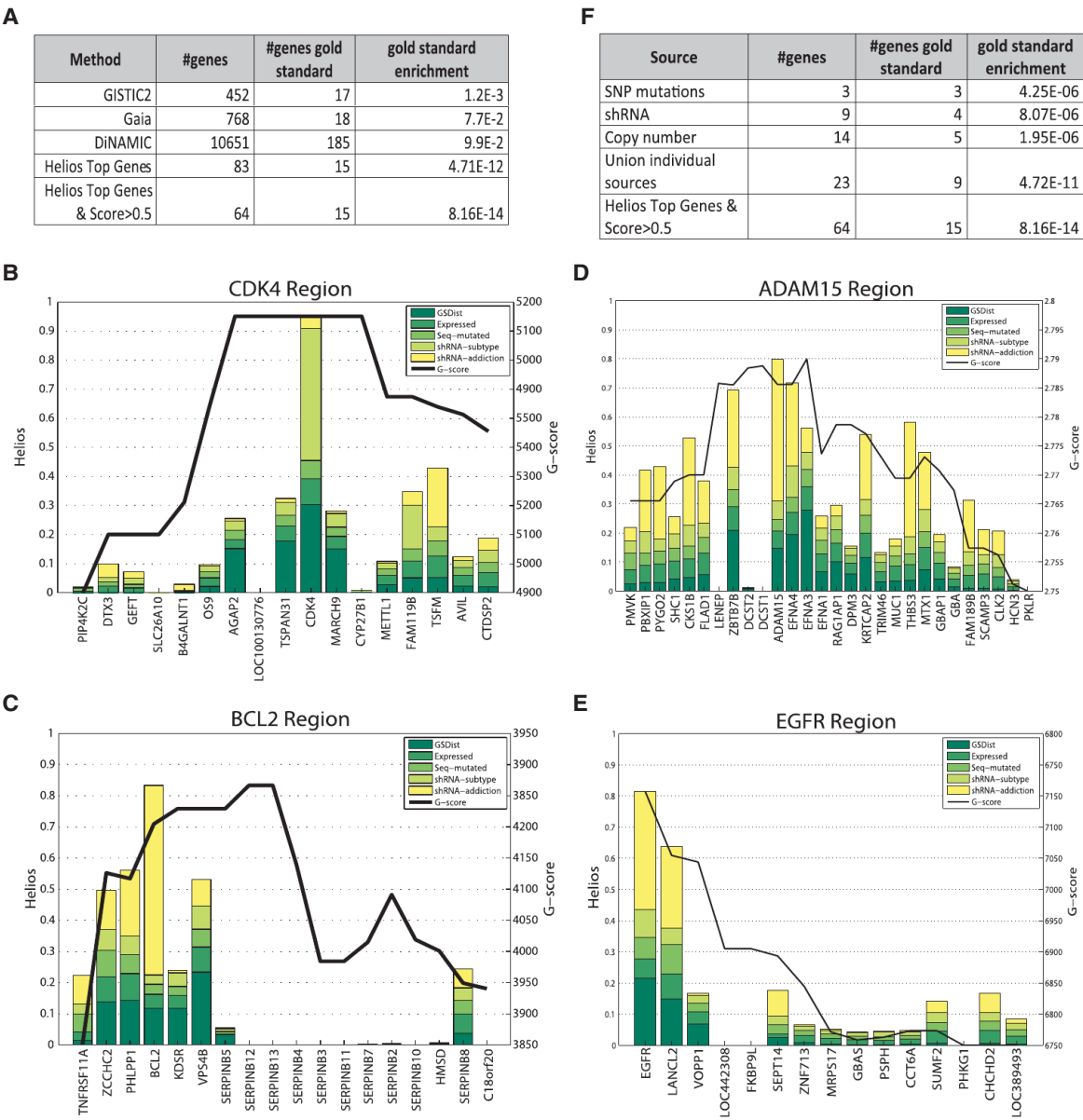


Figure 3. Helios Analysis of Breast Cancer

(A) A comparison of enrichment for a literature-compiled set of breast cancer drivers between our Helios genes, defined as the top gene in each region with a score >0.5 and three state of the art methods.

(B)–(E) display the result of the Helios analysis for the 12p14, 18q21, 1q21, and 7p12 regions, respectively. Genes in the ISAR regions are displayed in the x axis and the Helios score is represented by bars colored proportionally to the contribution of each feature (a logistic regression approximation is employed to approximate the contribution of each feature). The G score is displayed as a black line.

(F) A comparison between Helios and the results from the analysis of the data sources individually, testing for enrichment based on our literature compiled set of breast cancer drivers. See Figures S3A–S3D for information about convergence and stability of the results.

work that prioritized mostly kinases for experimental validation, for an unbiased evaluation of Helios, we chose a systematic score driven approach to validation. To perform an unbiased and comprehensive assessment, over a wide range of Helios scores, we used the independent ISAR score >5.5 to select regions and used Helios to pinpoint the most likely driver within each region. Thus, we sought to assess how often could Helios

pinpoint the correct driver for each of the 17 most frequently and significantly amplified regions.

In seven of the 17 regions, the top Helios gene was a bona-fide breast cancer oncogene (*ERBB2*, *CCND1*, *ZNF217*, *MYC*, *miR-21*, *FGFR2*, and *IGF1R*) and these oncogenes scored well above the next best scoring gene. For example, *MYC*'s Helios score was 100 times greater than the second best gene in the region

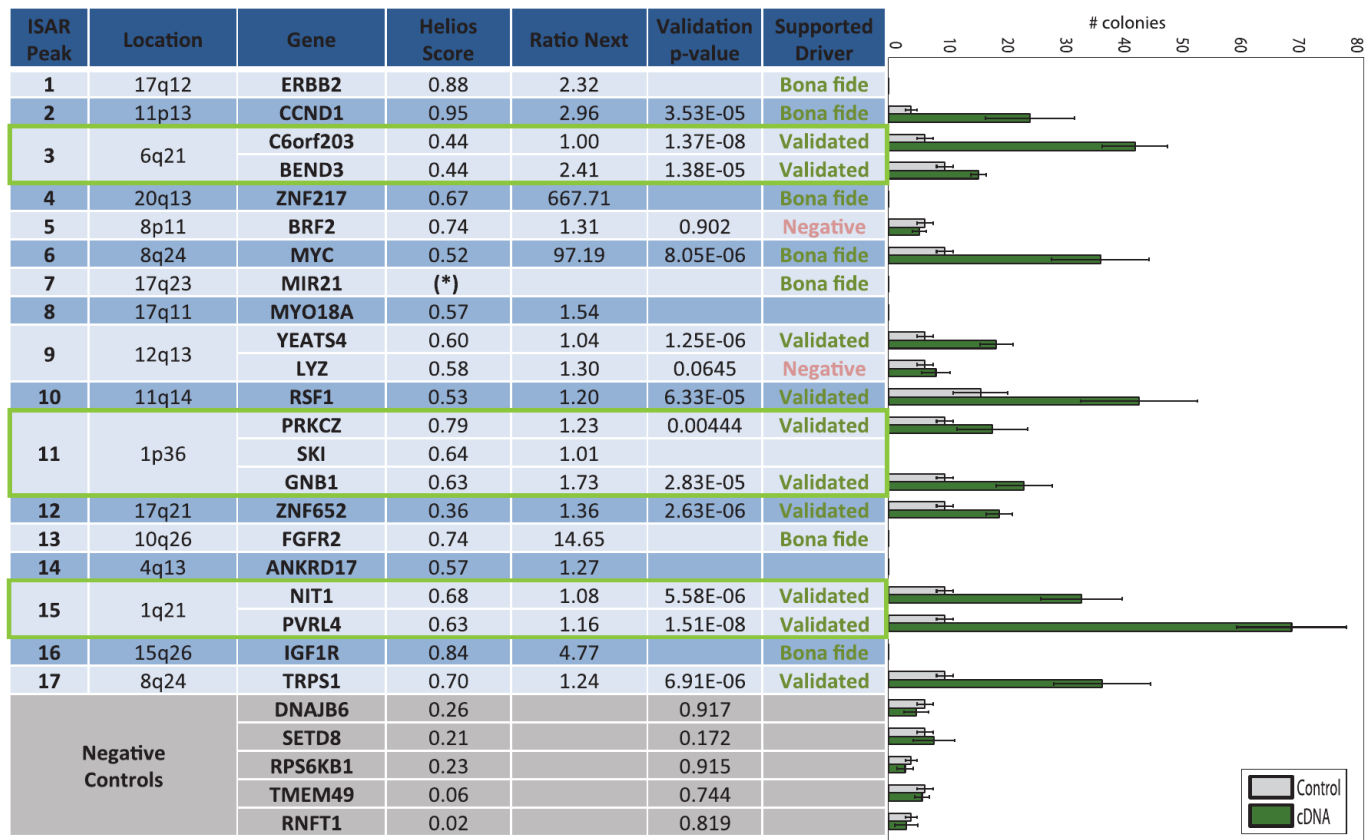


Figure 4. Helios Validations

Results of the systematic in vitro validation of Helios candidates including the selected genes and five genes selected as negative controls (highlighted in gray). One or more rows is shown for each of the top 17 highest scoring ISAR peaks regardless of whether or not any gene from the region was tested experimentally. The “Ratio Next” column indicates the ratio between the Helios score of the candidate gene and the score of the next best scoring gene in the region. The “Validation p value” column displays the statistical significance of the change in colony size between the six empty vector controls and the six repeats of the cDNA overexpressing the candidate driver gene. Error bars represent $2 \times$ SD of the replicates of the assay. The validation p value was computed using a right-tailed unpaired two-sample t test. The “Supported Driver” column indicates if the gene has been positively validated by the in vitro assay or is a known driver based on previous literature. The rightmost panel shows the box plots of the colony numbers for each gene in the validation experiment, where gray indicates the control and green the cDNA overexpressing the candidate driver gene. The colony assay was not performed for several genes that we failed to clone (MYO18A, SKI), or were bona fide drivers at the top of their peak (ERBB2, ZNF217, FGFR2, IGF1R). Additionally, no gene scored above 0.3 in the 17q23 region, suggesting that the target was another regulatory element, in this case the bona fide onco-microRNA MIR21. The three green boxes highlight amplified regions in which we confirmed more than one driver. The colony data that supports this figure is available in Figures S4B (candidate drivers) and S4C (negative controls). See also Table S2.

(Figure 4, “Ratio-next” column). There was no known breast cancer oncogene present among ten additional regions, and therefore we decided to perform in vitro validation for the top-scoring Helios genes in each of these regions. Because an amplified region can harbor more than one oncogene, we selected multiple genes if more than one scored significantly (4/10 regions). We failed to clone overexpression vectors for three genes, resulting in a final selection of 12 predicted oncogenes for validation. The selected genes encompassed a wide range of functional roles including chromatin remodeling, transcription factors, cell surface, cell adhesion proteins, and metabolic enzymes.

One of the hallmarks of transformation that is commonly used to investigate putative epithelial oncogenes is the ability to promote attachment-independent growth of a nontransformed cell line (Hanahan and Weinberg, 2011). This capacity likely reflects

the cumulative impact of multiple signals such as increased resistance to stress, increased cellular growth rates and changes in metabolism (Davison et al., 2013). As a result, many driver alterations in cancer may potentially impact attachment-independent growth through multiple mechanisms. Therefore, we based our candidate validation strategy on assaying this phenotype.

Experimental In Vitro Validation Confirms Helios-Predicted Genes

For each of the 12 candidate genes, we evaluated the ability of a clone of MCF-10A cells (human mammary epithelium) with intrinsic low attachment-independent growth ability, (see Experimental Procedures) to form colonies in semisolid media when the putative oncogene was experimentally upregulated. These cells were transduced with viral vectors overexpressing the

putative driver and evaluated for growth in soft agar. CCND1 and MYC were used as positive controls, and for negative controls, we selected five genes from significantly amplified ISAR regions (ISAR >5.5) that did not have a high Helios score (score <0.3). The agar assays for each gene was tested with a minimum of six replicates and statistical significance was evaluated by unpaired two-sample t test between the six test and six control plates.

Ten of 12 tested genes (*C6ORF23*, *BEND3*, *YEATS4*, *RSF-1*, *PRKCZ*, *GNB1*, *ZNF652*, *NIT1*, *PVRL4*, and *TRPS1*) were able to significantly increase MCF10A anchorage-independent activity with a p value of 0.005 or below (Figures 4 and S4). None of the negative controls demonstrated an increase in colony formation. This provides in vitro evidence that Helios is highly specific in identifying genes that provide a selective advantage for breast cancer cells. Note that a negative result for BRF2 (demonstrated to be an oncogene in lung cancer) (Lockwood et al., 2010) does not conclusively rule it out as a driver gene, because attachment-independent growth is not the only hallmark of cancer and the assays were performed in a single genetic background.

Overall, Helios demonstrated unprecedented accuracy in identifying genes that promote oncogenic capabilities. Helios correctly scored 13/14 drivers at the top of their respected region (93%). Moreover, 10/12 empirically tested genes validated (83%), thus we identified ten genes that promote tumorigenic capabilities in breast cancer (including *PVRL4* that was recently published [Pavlova et al., 2013]). Additionally, because the genes were selected based on the region's significance, rather than their Helios score, a wide range of Helios scores were tested (between 0.36 and 0.79), increasing our confidence in the candidates identified in other regions. Based on this performance, we expanded our list of likely drivers based on Helios predictions with more permissive criteria (Table S3).

Importantly, Helios identified multiple high scoring (likelihood >0.5) genes for over 20% of the regions. Indeed, we validated three regions with multiple genes and each gene independently induced colony formation in vitro (Figure 4, green boxes), indicating that an amplicon often targets more than one gene. In summary, while previously only 7/17 of the most frequently altered regions in breast cancer harbored a known oncogene, following our validation 14/17 regions can be assigned a driver with substantial confidence.

RSF-1 Promotes Colony Growth In Vitro

Among the ten validated candidates, RSF-1 is an especially compelling putative driver because it is recurrently amplified in several cancers (Chen et al., 2011; Fang et al., 2011; Li et al., 2012; Liu et al., 2012; Shih et al., 2005). Additionally, an amplicon containing RSF-1 was recently associated with a breast cancer subtype bearing one of the worst clinical prognoses (Curtis et al., 2012). Although high expression levels of *RSF-1* has been associated with poor prognosis in several malignancies (Hu et al., 2012; Li et al., 2012; Liu et al., 2012; Sheu et al., 2013), its involvement in breast cancer pathogenesis has not yet been explicitly demonstrated. Therefore, we chose to follow-up our analysis of *RSF-1* with further in vitro and in vivo experiments.

We selected four additional mammary epithelial cell lines nonamplified for RSF-1. The human MCF-10A-Triple Modified (a MCF-10A variant sensitized to transformation called here

MCF-10A-TM) (Pires et al., 2013), MDA-MB-415, MDA-MB-361, and the mouse Comma-ID (C-ID) (Campbell et al., 1988). We also selected one cell line (MDA-MB-453, human) with amplified and overexpressed RSF-1 (Figure S5A). Overexpression of *RSF-1* in all nonamplified cell lines increased the ability to form colonies in semisolid media (Figure 5A). To assay RSF-1 oncogene addiction, we selected two doxycycline (Dox) inducible shRNA-miRs that efficiently silenced *RSF-1* and assayed colony formation of the RSF-1 amplified MDA-MB-453 line. As expected, silencing of *RSF-1* significantly reduced the number of colonies formed (Figure 5B). To demonstrate that the loss of tumorigenic potential is not an off-target effect, we restored RSF1 expression in these cells by overexpressing the RSF1 cDNA (Figure S5C). Restoring RSF1 levels rescued the ability of MDA-MB-453 to form colonies in agar despite the expression of RSF1 shRNAs.

RSF-1 Promotes Growth in Xenograft Models

Next, we conducted experiments to assay RSF-1 in vivo. MCF-10A-TM and C-ID were orthotopically transplanted into the fat pad of severe combined immunodeficiency (SCID) mice with and without prior transduction of an RSF-1 overexpression vector. We then tracked the development of tumors and compared growth between controls and those overexpressing *RSF-1*.

MCF-10A cells are not tumorigenic, and overexpression of *RSF-1* did not transform them. While some transplanted MCF-10A-TM cells remained in the fat pad, these did not produce tumor. However, MCF-10A-TM overexpressing *RSF-1* was able to establish small primary tumor outgrowths (Figures 5C and S5B). C-ID overexpressing *RSF-1* cells generated palpable masses as early as 2 weeks after transplantation—significantly earlier than control mice, which lacked detectable tumor burden after 1 month. ($p = 0.0001$) (Figures 5D and S5B).

Finally, we also transplanted *RSF-1* amplified MDA-MB-453 cells and an MDA-MB-453 variant bearing a doxycycline inducible RSF-1-ShRNA into the fat pad of SCID mice. As expected, in the absence of Dox, all MDA-MB-453 variants generated tumors that grew at a comparable rate. However, supplementing the mice with Dox reduced the tumorigenic growth specifically in the tumors carrying the RSF-1 shRNA (Figure 5E). This data provides evidence that RSF-1 can contribute to tumor progression in vivo and that inhibition of RSF-1 expression can cause tumor regression.

RSF-1 Promotes Invasion in Xenograft Models

To further characterize the role of RSF-1 in breast cancer, we analyzed the TCGA gene expression data and identified gene-expression signatures associated with RSF-1 expression levels (Akavia et al., 2010; Danussi et al., 2013) (Extended Experimental Procedures). Genes associated with RSF-1 in this procedure are putative downstream targets of RSF-1 activity. We performed gene set enrichment in these signatures using the MSigDB database (Subramanian et al., 2005) and found enrichment for gene sets involved in invasion, metastasis, and de-differentiation (Figures 6A and S6A).

Therefore, we hypothesized that *RSF-1* overexpression may promote metastatic potential in vivo. To test this, we performed intravenous tail injection of MCF-10A-TM cells expressing a

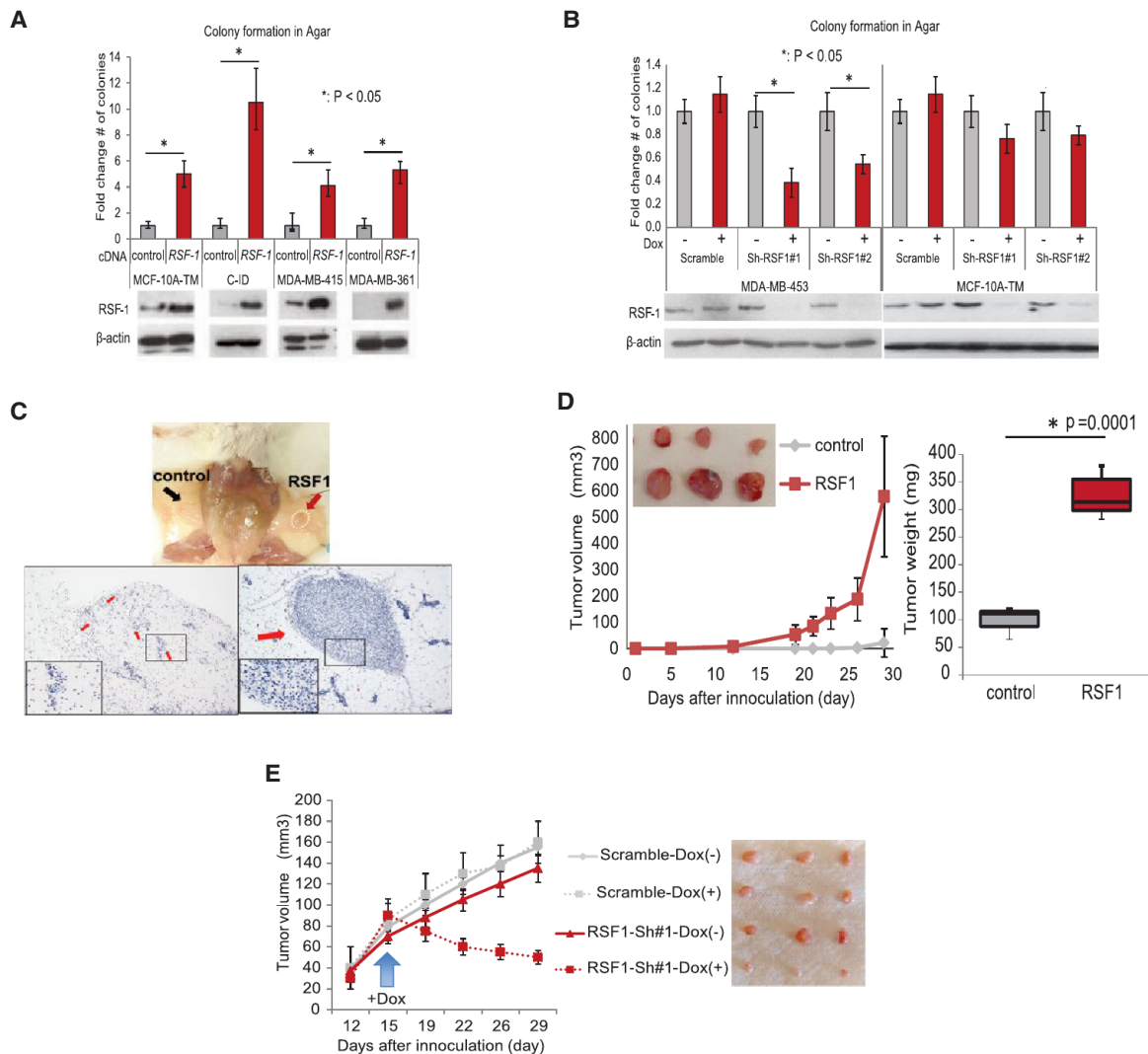


Figure 5. High Expression Levels of RSF-1 Promote Tumorigenesis

(A) Overexpression of RSF-1 in multiple cell lines enhances its ability to form colonies in agar.

(B) downregulation of RSF-1 using Dox-inducible shRNAs in a cell line with amplification of the locus (MDA-MB-453) reduced its ability to form colonies in agar.

(C and D) Overexpression of RSF-1 in (C) MCF-10A-TM and (D) CID cells enhanced their tumorigenic potential in vivo. (C) Overexpression of RSF1 in MCF-10A-TM cells resulted in the formation of small tumor masses demonstrated in hematoxylin and eosin (H&E) images at right compared to control at left. Number of tumors formed for each model is available in Figure S5B. (D) Overexpression of RSF1 in CID cells increased tumor volume (left) and weight (right) following orthotopic transplantation in SCID mice.

(E) Silencing of RSF-1 in MDA-MB-453 attenuated its tumorigenic potential when orthotopically transplanted in SCID mice. Error bars represent SD of $n = 3$ for in vitro and $n = 6-8$ for in vivo studies. Animal studies comply with IACUC regulations.

luciferase reporter into SCID mice. When cells are injected intravenously in the tail of recipient mice, the cells travel through the circulatory system and are deposited in the lungs, where the majority of the cells die due to the absence of a supportive micro-environment (Yang et al., 2012). Both control and *RSF-1* overexpressing cells were rapidly cleared and no signal was detected 1 week after the injection. Importantly, after 7 weeks, all the mice injected with cells overexpressing *RSF-1* showed luciferase signal in the lungs indicating the formation of lung metastases while luciferase signal was never recovered in mice injected with control cells (Figure 6B). This demonstrates that RSF-1 over-

expression promotes increased invasive capacity in the lungs and therefore a prometastatic state in breast cancer cells.

In summary, we have shown that overexpression of RSF-1 confers increased anchorage-independent growth in vitro and promotes the formation of lung metastases in mouse models. Additionally, we have identified a transcriptional signature associated with RSF-1 amplification in primary tumors that was enriched for genes related to metastasis and invasion. The identification of RSF-1 as an oncogene that increases metastatic potential provides an explanation for the steep mortality of a recently identified molecular subgroup of breast cancer (Curtis et al., 2012).

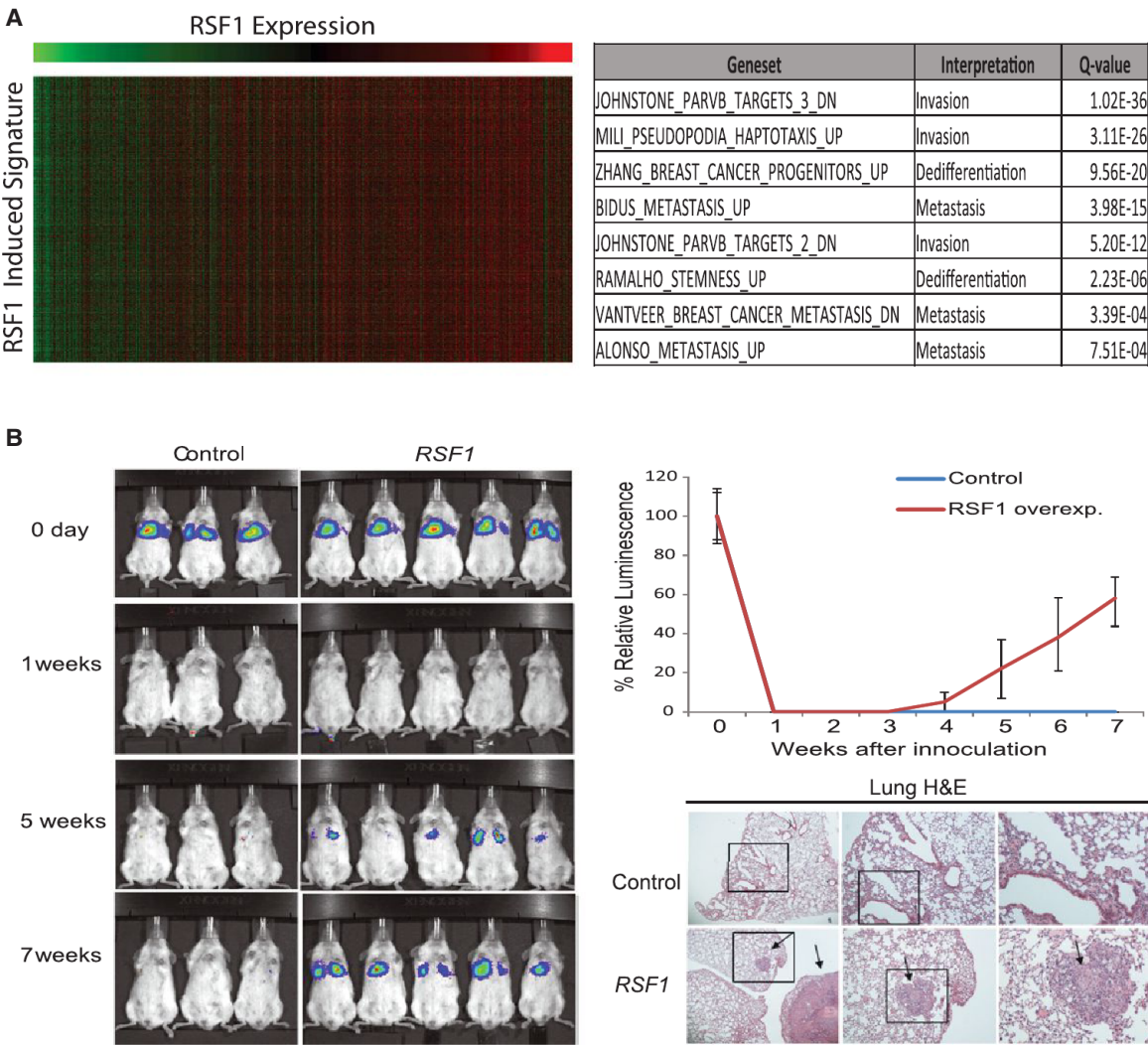


Figure 6. RSF-1 Alteration Promotes Metastasis
(A) The analysis of the expression changes related to RSF-1 overexpression in basal primary tumors revealed a signature enriched for invasiveness, migration and dedifferentiation (table at right). The heat map at left shows genes in the signature as rows and samples as columns and the color indicates the relative expression (green-low and red-high) and demonstrates the tight correlation of the signature genes across patients. Similar results were observed for luminal primary tumors (Figure S6A). See Figures S6B and S6C for analysis of downregulated genes.
(B) Comparison of lung metastasis formation in SCID mice subjected to tail vein injection of MCF-10A-TM cells expressing a luciferase reporter and either an RSF-1 overexpression vector or a control vector. H&E of sectioned lungs from mice injected with control and RSF-1 overexpressing cells is also shown. The arrows indicate the presence of metastatic outgrowths in the lungs. Animal studies comply with IACUC regulations.

DISCUSSION

Cancer research has recently been driven by the hope that therapies targeting drivers will be especially effective in tumors harboring genetic alterations in the target. This approach relies on the oncogene addiction effect whereby cancer cells become dependent on the activity of their altered oncogenes, so that inhibiting them compromises cellular viability. This “personalized medicine” is the basis of some of the most effective therapies, e.g., those targeting ERBB2 amplification in breast cancer (Ashworth et al., 2011). The success of these therapies has fueled efforts to catalog the genomic alterations in numerous cancers

with the hopes of discovering new therapeutically actionable mutations.
However, even as data from cancer genomes accumulates, the identification of actionable driver genes remains a crucial limitation to therapeutic development. We see at least two significant bottlenecks. First, only a small subset of established driver genes are druggable given the current pharmacological state of the art (Collins and Workman, 2006). Second, even when a driver is druggable, it may occur in a very small fraction of patients, limiting its clinical utility. At present, there is an untapped resource of driver genes in SCNAs that have evaded discovery. Moreover, due to the high frequency of SCNA events, actionable

drivers can impact more patients (Figure 1A). However, to date, this possibility has been limited by the difficulty of distinguishing passengers and drivers in the majority of SCNAs.

Here, we have presented a major advance in addressing this challenge, using a method that integrates data from primary tumors with functional assays on cell lines to prioritize candidate drivers. The unparalleled sensitivity and specificity of Helios enabled us to execute the first reported systematic validation of an algorithm designed to identify driver genes. Helios's performance was confirmed by a success rate of 10/12 candidates in an anchorage-independent growth assay, successfully characterizing several regions for which there was no previously implicated driver. Importantly, because we selected the genes for validation based on their amplification significance (ISAR score), rather than their Helios score, we expect that this success rate will extend to additional regions that have equally strong Helios scores. Moreover, many of these genes are amplified in additional epithelial cancers (e.g., *C6orf203*, *NIT1*, *ZNF652*) suggesting possible drivers in those cancers as well.

Using Helios, we have significantly expanded the landscape of high-confidence breast cancer drivers by more than 2-fold (Figures 7 and S7). Previous analyses of breast cancer cohorts (Stephens et al., 2012; Cancer Genome Atlas Network, 2012) had identified 15 driver genes occurring in at least 5% of breast cancer tumors (both SCNA and sequence mutations). Our analysis has doubled this number to 29, substantially expanding the list of potential drug targets. Even more importantly, we have increased the number of drivers identified in each tumor, thus raising the possibility that at least one might be actionable in a given patient. A previous study (Figure 7B, gray boxes) (Stephens et al., 2012) could assign each tumor a median of two established drivers. Adding the Helios validated genes increases this number to a median of three drivers per tumor (Figure 7B, green boxes). Adding all predicted drivers with a high Helios score further expands this number to a median of five drivers in each tumor (Figure 7B, yellow boxes). Thus Helios has substantially expanded the set of high-confidence drivers in breast cancer.

Helios uses a technique called transfer learning, whereby drivers with clearer signal (e.g., at the peak of their region) help learn informative features to improve performance in cases with less obvious signal. Helios learns the list of candidate drivers without using any prior list of driver genes and therefore it does not suffer from any bias. The algorithm uses all data in its learning process, transferring information across different genes, as well as between SCNAs and other features, until it converges into a final ranking of candidate driver genes. By leveraging information in this fashion, Helios is capable of learning how to weigh and combine features into a probabilistic score that represents the likelihood of the gene being the target of the recurrent alteration. This computational framework is independent of the features and tumor type and it can be applied to analyze additional cancers using a similar or even different set of features.

Genetic, genomic and functional data on cancers will continue to accumulate from large-scale projects in the coming years (Cheung et al., 2011; Cancer Genome Atlas Network, 2008). Such data sets continue to accelerate drug development and

to yield deep insights into oncogenesis. However, they also create new analytical challenges such as the need to pinpoint the alterations that promote cancer. Helios can be viewed as an accurate in silico screen for drivers. As such, it can be applied to additional cancer types and data types to accelerate the identification of cancer drivers.

EXPERIMENTAL PROCEDURES

ISAR

ISAR is based on the G score metric, a significance measure of the aberration for each marker, which was originally defined in GISTIC (Beroukhi et al., 2007). Specifically, the G score for a marker m is the summation of the copy number across samples that surpass an aberration threshold θ . Given the copy number for n samples, the G score for a marker m in the case of amplifications is:

$$G^{AMP}(m) = \sum_{i=1}^N CN(m, i) \times I(CN(m, i) > \theta^{AMP}), \quad (\text{Equation 1})$$

where $CN(m, i)$ is the copy number of marker m in sample i and I is the indicator function.

ISAR uses a local sliding window of constant size that moves along the chromosome, calculating the null distribution for each window. Once the distribution has been computed in all windows within a chromosome, each genomic marker is associated with several overlapping windows. The algorithm takes a conservative approach by selecting the least significant q value among the values computed for all overlapping windows containing the marker (see Extended Experimental Procedures for more detail).

Modeling Copy Number

We aim to model a distribution of SCNA that reflects the differences between driver and passenger genes, independently of the chromosomal region. However, in contrast to the subtle differences in SCNA within each altered region, the distribution of alterations differs dramatically between regions. Indeed, the median difference in G score between genes in a region is significantly smaller (172) than the difference for genes across different regions (6,405). Thus, without appropriate normalization, the G score should not be used to prioritize drivers across regions. We aim to model whether the gene is among the most altered genes in its own region (and therefore more likely to be the driver of that region) and therefore define a metric that measures the difference in terms of G score to the highest value in each region. For a single gene g , we define the GSDist score as:

$$GSDist(g) = \max_{j \in \text{region}(g)} (Gscore(j) - Gscore(g)). \quad (\text{Equation 2})$$

The most altered gene(s) in a region will have $GSDist = 0$, while any other gene will have a positive value that indicates the "distance" to the most frequently amplified gene in the region. Note that traditional approaches would use a threshold on this metric to make a hard decision on whether genes in the altered region are peak genes (Figure 2B). Instead Helios models this metric using two exponential distributions (one for drivers and one for passengers):

$$P(SCNA|\lambda_t) = \lambda_t e^{-\lambda_t GSDist}. \quad (\text{Equation 3})$$

Driver genes have a GSDist distribution that exponentially decreases from zero with small variance, whereas passenger genes are modeled by a uniform distribution, which is approximated by an exponential distribution with large variance (see Extended Experimental Procedures for more detail).

Features Used in the Helios Algorithm

We use MutSig (Banerji et al., 2012) to compute the statistical significance of the recurrence of point mutations.

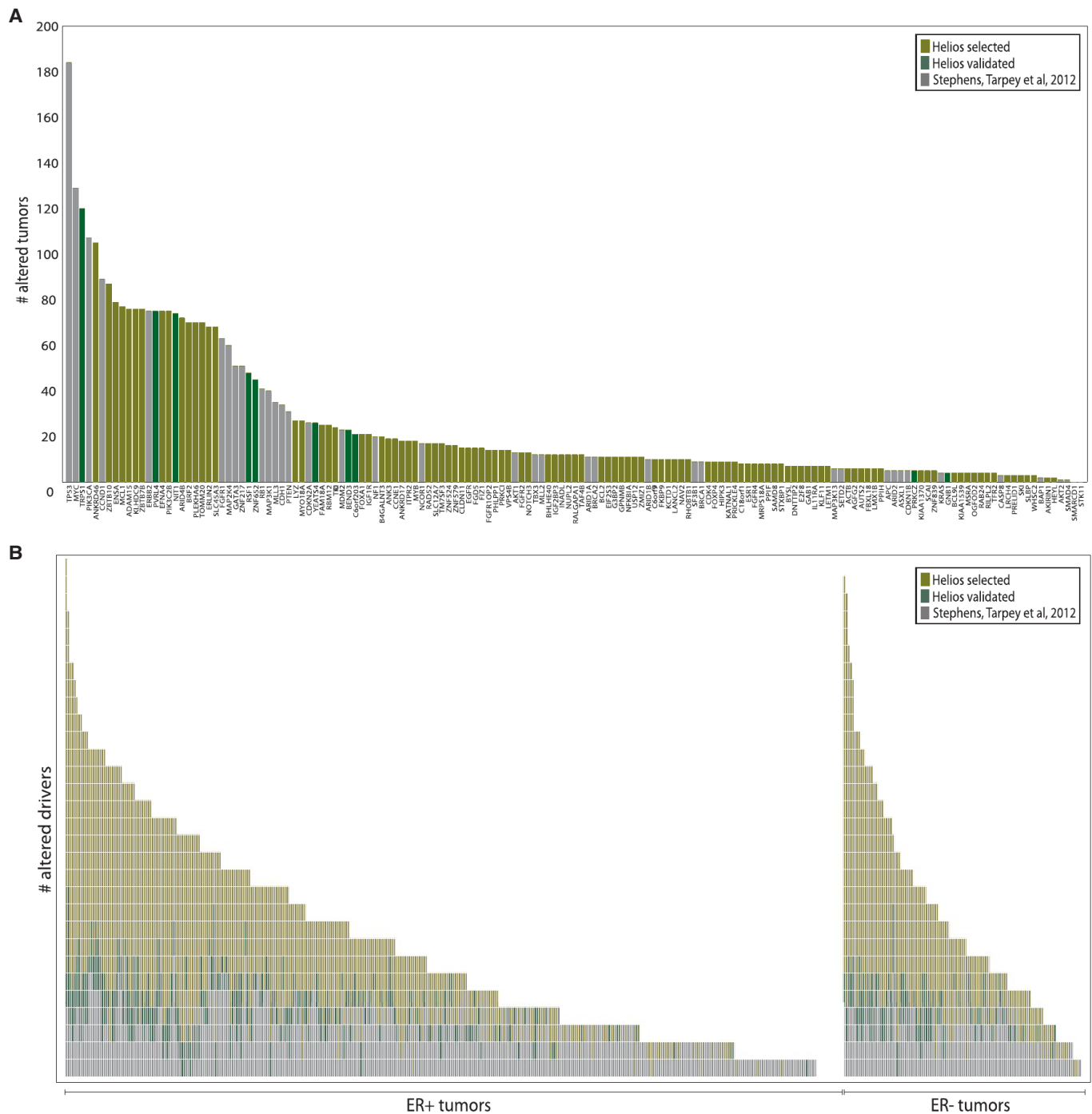


Figure 7. The Landscape of Driver Mutations in Breast Cancer

(A and B) For the driver genes described in Stephens et al. (2012) (gray), Helios validated genes (green), and other Helios genes scoring >5.5 (yellow), we compute (A) the number of tumors altered (copy number or sequence mutation) for each driver gene and (B) the number of driver genes altered (copy number or sequence mutation) per tumor. For this figure we consider the 485 primary tumors in TCGA for which both copy number and DNA-Seq were available.

See also Figure S7 and Table S3.

Helios uses features extracted from RNA-Seq-based gene expression in two different ways: (1) to identify genes that are not expressed and are therefore unlikely to be drivers, and (2) we expect the oncogenic activity of an amplified driver gene to be reflected in the gene's mRNA dosage (Akavia et al., 2010).

The oncogene addiction score for a hairpin is defined as the log-likelihood of the monotonic regression that predicts the lethality based on the gene mRNA. We use the PAVA algorithm (Brunk, 1955) to estimate the best fit for the regression (see Extended Experimental Procedures for more detail).

Helios Algorithm

Helios uses a hierarchical Bayesian mixture model to distinguish drivers from passengers among the genes present in significantly altered regions. The unsupervised Bayesian algorithm discriminates driver genes ($T = 1$) by integrating the copy-number alteration information (SCNA), with cues from different data sources (X). The hierarchical framework naturally separates these two components using the following model:

$$P(CNA) = \sum_{t \in \{0,1\}} P(SCNA|T=t)P(T=t|X). \quad (\text{Equation 4})$$

This model separates the modeling of copy number ($P(SCNA|T=t)$) from other sources of information ($P(T=t|X)$), focusing on predicting the observed copy-number landscape ($P(SCNA)$). The algorithm iteratively fits a model for each part: $P(SCNA|T=t)$ and $P(T=t|X)$ and updates the estimations for each gene (T) taking both parts into account. The algorithm is executed until the model converges into a stable solution that incorporates all the information into a single probability score for each gene.

Figure S2A shows the graphical model for Helios, where n genes are classified by combining the information from different data sources X and SCNA. w represents the parameters that control the integration of X , while λ parameterizes the influence of SCNA. In this model, when the values T_n for the genes are given, the parameters for the different sources (w) and copy number (λ) are independent. This property makes it possible to fit the model efficiently using the Expectation Maximization (EM) algorithm (see Extended Experimental Procedures for more details).

Data Sets Used for Helios

We used the following public data sets:

Primary tumor data from the TCGA project (Cancer Genome Atlas Network, 2012): copy number Affymetrix 6.0 SNP arrays ($n = 785$), Illumina HiSeq RNA sequencing ($n = 732$), and whole-exome sequencing ($n = 507$).

Cell line shRNA screens ($n = 29$) collected by Marcotte et al. (2012).

Cell line data from the Cancer Cell Line Encyclopedia (Barretina et al., 2012) for the cell lines screened with shRNA: copy number Affymetrix 6.0 SNP arrays ($n = 27$) and messenger RNA Affymetrix U133 plus 2.0 arrays ($n = 27$).

Data Sets Used to Generate Gold Standard Set

To assess performance, a gold standard set of 330 genes was compiled from the following sources:

The set of known amplified oncogenes from Beroukhi et al. (2010).

The set of genes related to breast cancer according to the University of Copenhagen DISEASES database (Pletscher-Frankild et al., 2014) with score >2.5 . We filtered out genes categorized as tumor suppressors according to the Uniprot Consortium (2013).

See Extended Experimental Procedures for more information.

Cell Culture and Reagents

To generate cell lines overexpressing a gene, cells were plated at 60% of confluence in a 6 well plate and after 24 hr infected with virus expressing the different plasmids containing the different genes. Media containing virus was replaced in 12 hr for fresh media. Cells were then reinfected for another 12 hr. Cells were grown in fresh media for 24 hr and selected with the appropriate drug. To generate MDA-MB 453 cells deficient in *RSF1* expression, cells were infected with virus expressing doxycycline-inducible pTRIPz shRNAs against *RSF1*. Then, cells expressing the shRNAs were selected with the puromycin (2 $\mu\text{g/ml}$).

See Extended Experimental Procedures for cell lines, DNA constructs, and gene cloning strategy.

Validation of Helios predictions was based on the ability of MCF-10A to form colonies in semisolid media when the putative oncogene was experimentally upregulated. Because low passage MCF-10A are very resistant to transformation, to increase the sensitivity of our assay, we selected a passage with intrinsic low attachment-independent growth ability (5–15 colonies per 5,000 plated cells) that demonstrated robust higher growth ability when bona-fide breast oncogenes were overexpressed (Figure S4A).

Colony formation assay in semisolid media was performed in 6 well plates. First, a layer of 2 ml of 0.6% agar (Fisher 9002-18-0) in regular MCF-10A media

was placed at the bottom of each well and allowed gelification. Then, a layer of 2 ml of 0.3% agar containing 5,000 cells was seeded on top of the bottom agar layer and allowed gelification. Finally, 1 ml of regular MCF-10A media was placed covering the agar. The colonies were allowed to form for 1 month. After this period, 2 ml of MTT solution (Sigma M5655) at 0.5 mg/ml was used to stain the colonies. A minimum of six replicas per gene were plated. To ensure comparability, transformation assays for each gene are compared to empty-vector controls performed together on the same day. The number of colonies was independently evaluated by two researchers. All the different MCF-10A clones carrying controls and genes of interest were maintained growing exponentially for 48 hr (plates were at 50–70 confluence) before being plated in agar to homogenize assay conditions.

Tumorigenicity in Mice

Animal maintenance and experiments were performed in accordance with the animal care guidelines and protocols approved by Columbia University animal care unit. For Comma-1D cell line, 21-day-old female NOD SCID immunocompromised mice NOD.CB17-Prkdc SCID mice (Harlan) mice were injected with 5×10^5 cells, resuspended in PBS into a fat mammary gland. For MDA-453 cell line, 8-week-old female NOD SCID immunocompromised mice NOD.CB17-Prkdc SCID mice (Harlan) mice were injected with 5×10^6 cells, resuspended in 1:2 Matrigel (BD Biosciences) plus normal growth media into a fat pad mammary gland. Doxycycline was added to drinking water at a final concentration of 2.0 mg/ml. Tumor growth was monitored twice a week with calipers at the site of injection. Animals were sacrificed as soon as tumor size reached 1.5 cm in diameter.

In the experimental metastasis assays, 8-week-old female NOD SCID immunocompromised NOD.CB17-Prkdc SCID mice (Harlan) where injected with 5×10^6 cells, resuspended in PBS, via the tail vein. To measure the luciferase intensity of injected cells, 2.25 $\mu\text{g ml}^{-1}$ luciferin was injected intravenously through the tail and luciferase activity was assessed 5 min after luciferin injection using a IVIS Spectrum Pre-Clinical In Vivo Imaging System (IVISPE; Perkin-Elmer) machine. The presence of established metastases was confirmed by euthanizing the mice.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and five tables and can be found online at <http://dx.doi.org/10.1016/j.cell.2014.10.048>.

AUTHOR CONTRIBUTIONS

F.S.G., J.M.S., and D.P. conceived the study. F.S.G. and D.P. designed ISAR and HELIOS. F.S.G. designed and implemented ISAR and HELIOS. F.S.G., D.K., and B.J. preprocessed data. F.S.G., D.K., U.D.A., and D.P. performed statistical analysis of Helios results. F.S.G., P.V., D.K., J.M.S., and D.P. performed biological interpretation of Helios and selected genes for validation. P.V. and J.M.S. designed biological validation experiments. P.V., J.M., V.C., L.S.C., R.D.B., D.L.N., and J.M.S. performed biological validation of Helios genes. P.V., J.M., V.C., and J.M.S. performed characterization experiments for *RSF1*. F.S.G., P.V., D.K., J.M.S., and D.P. wrote the manuscript.

ACKNOWLEDGMENTS

The authors would like to thank Sourav Bandyopadhyay, Oren Litvin, Richard Marcotte, Ben Neel, Ramon Parsons, and Sagi Shapira for valuable comments. This research was supported by the NIH (grants R01CA164729 and DP2OD002414) and a National Centers for Biomedical Computing grant (1U54CA121852-01A1). D.P. holds a Packard Fellowship for Science and Engineering.

Received: January 21, 2014

Revised: August 1, 2014

Accepted: October 3, 2014

Published: November 26, 2014

REFERENCES

- Akavia, U.D., Litvin, O., Kim, J., Sanchez-Garcia, F., Kotliar, D., Causton, H.C., Pochanard, P., Mozes, E., Garraway, L.A., and Pe'er, D. (2010). An integrated approach to uncover drivers of cancer. *Cell* 143, 1005–1017.
- Albertson, D.G., Collins, C., McCormick, F., and Gray, J.W. (2003). Chromosome aberrations in solid tumors. *Nat. Genet.* 34, 369–376.
- Ashworth, A., Lord, C.J., and Reis-Filho, J.S. (2011). Genetic interactions in cancer progression and treatment. *Cell* 145, 30–38.
- Banerji, S., Cibulskis, K., Rangel-Escareno, C., Brown, K.K., Carter, S.L., Frederick, A.M., Lawrence, M.S., Sivachenko, A.Y., Sougnez, C., Zou, L., et al. (2012). Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature* 486, 405–409.
- Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G.V., Sonkin, D., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 483, 603–607.
- Beroukhi, R., Getz, G., Nghiemphu, L., Barretina, J., Hsueh, T., Linhart, D., Vivanco, I., Lee, J.C., Huang, J.H., Alexander, S., et al. (2007). Assessing the significance of chromosomal aberrations in cancer: Methodology and application to glioma. *Proc. Natl. Acad. Sci.* 104, 20007–20012.
- Beroukhi, R., Mermel, C.H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., Barretina, J., Boehm, J.S., Dobson, J., Urashima, M., et al. (2010). The landscape of somatic copy-number alteration across human cancers. *Nature* 463, 899–905.
- Bissell, M.J., and Hines, W.C. (2011). Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. *Nat. Med.* 17, 320–329.
- Brunk, H.D. (1955). Maximum likelihood estimates of monotone parameters. *Ann. Math. Stat.* 26, 607–616.
- Campbell, S.M., Taha, M.M., Medina, D., and Rosen, J.M. (1988). A clonal derivative of mammary epithelial cell line COMMA-D retains stem cell characteristics of unique morphological and functional heterogeneity. *Exp. Cell Res.* 177, 109–121.
- Cancer Genome Atlas Network (2008). Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 455, 1061–1068.
- Cancer Genome Atlas Network (2012). Comprehensive molecular portraits of human breast tumours. *Nature* 490, 61–70.
- Chen, T.J., Huang, S.C., Huang, H.Y., Wei, Y.C., and Li, C.F. (2011). Rsf-1/HBXAP overexpression is associated with disease-specific survival of patients with gallbladder carcinoma. *APMIS: acta pathologica, microbiologica, et immunologica Scandinavica* 119, 808–814.
- Cheung, H.W., Cowley, G.S., Weir, B.A., Boehm, J.S., Rusin, S., Scott, J.A., East, A., Ali, L.D., Lizotte, P.H., Wong, T.C., et al. (2011). Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer. *Proc. Natl. Acad. Sci. USA* 108, 12372–12377.
- Collins, I., and Workman, P. (2006). New approaches to molecular cancer therapeutics. *Nat. Chem. Biol.* 2, 689–700.
- Curtis, C., Shah, S.P., Chin, S.-F., Turashvili, G., Rueda, O.M., Dunning, M.J., Speed, D., Lynch, A.G., Samarajiwa, S., Yuan, Y., et al.; METABRIC Group (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* 486, 346–352.
- Danussi, C., Akavia, U.D., Niola, F., Jovic, A., Lasorella, A., Pe'er, D., and Iavarone, A. (2013). RHPN2 Drives Mesenchymal Transformation in Malignant Glioma by Triggering RhoA Activation. *Cancer research*.
- Davison, C.A., Durbin, S.M., Thau, M.R., Zellmer, V.R., Chapman, S.E., Diener, J., Wathen, C., Leevy, W.M., and Schafer, Z.T. (2013). Antioxidant enzymes mediate survival of breast cancer cells deprived of extracellular matrix. *Cancer Res.* 73, 3704–3715.
- De, S., and Michor, F. (2011). DNA secondary structures and epigenetic determinants of cancer genome evolution. *Nat. Struct. Mol. Biol.* 18, 950–955.
- Fang, F.M., Li, C.F., Huang, H.Y., Lai, M.T., Chen, C.M., Chiu, I.W., Wang, T.L., Tsai, F.J., Shih, I.M., and Sheu, J.J. (2011). Overexpression of a chromatin remodeling factor, RSF-1/HBXAP, correlates with aggressive oral squamous cell carcinoma. *Am. J. Pathol.* 178, 2407–2415.
- Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674.
- Hu, B.S., Yu, H.F., Zhao, G., and Zha, T.Z. (2012). High RSF-1 expression correlates with poor prognosis in patients with gastric adenocarcinoma. *Int. J. Clin. Exp. Pathol.* 5, 668–673.
- Kaelin, W.G., Jr. (2012). Molecular biology. Use and abuse of RNAi to study mammalian gene function. *Science* 337, 421–422.
- Kandoth, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M.A., et al. (2013). Mutational landscape and significance across 12 major cancer types. *Nature* 502, 333–339.
- Li, Q., Dong, Q., and Wang, E. (2012). Rsf-1 is overexpressed in non-small cell lung cancers and regulates cyclinD1 expression and ERK activity. *Biochem. Biophys. Res. Commun.* 420, 6–10.
- Liu, S., Dong, Q., and Wang, E. (2012). Rsf-1 overexpression correlates with poor prognosis and cell proliferation in colon cancer. *Tumour biology* 33, 1485–1491.
- Lockwood, W.W., Chari, R., Coe, B.P., Thu, K.L., Garnis, C., Malloff, C.A., Campbell, J., Williams, A.C., Hwang, D., Zhu, C.-Q., et al. (2010). Integrative genomic analyses identify BRF2 as a novel lineage-specific oncogene in lung squamous cell carcinoma. *PLoS Med.* 7, e1000315.
- Lohr, J.G., Stojanov, P., Lawrence, M.S., Auclair, D., Chapuy, B., Sougnez, C., Cruz-Gordillo, P., Knoechel, B., Asmann, Y.W., Slager, S.L., et al. (2012). Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc. Natl. Acad. Sci. USA* 109, 3879–3884.
- Marcotte, R., Brown, K.R., Suarez, F., Sayad, A., Karamboulas, K., Krzyzanowski, P.M., Sircoulomb, F., Medrano, M., Fedyszyn, Y., Koh, J.L.Y., et al. (2012). Essential gene profiles in breast, pancreatic, and ovarian cancer cells. *Cancer Discov.* 2, 172–189.
- Mermel, C.H., Schumacher, S.E., Hill, B., Meyerson, M.L., Beroukhi, R., and Getz, G. (2011). GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* 12, R41.
- Morganella, S., Pagnotta, S.M., and Ceccarelli, M. (2011). Finding recurrent copy number alterations preserving within-sample homogeneity. *Bioinformatics* 27, 2949–2956.
- O'Day, E., and Lal, A. (2010). MicroRNAs and their target gene networks in breast cancer. *Breast Cancer Res.* 12, 201.
- Pavlova, N.N., Pallasch, C., Elia, A.E., Braun, C.J., Westbrook, T.F., Hemann, M., Elledge, S.J., and Staudt, L. (2013). A role for PVRL4-driven cell-cell interactions in tumorigenesis. *eLife* 2, e00358.
- Pires, M.M., Hopkins, B.D., Saal, L.H., and Parsons, R.E. (2013). Alterations of EGFR, p53 and PTEN that mimic changes found in basal-like breast cancer promote transformation of human mammary epithelial cells. *Cancer Biol. Ther.* 14, 246–253.
- Pletscher-Frankild, S., Pallegà, A., Tsafou, K., Binder, J.X., and Jensen, L.J. (2014). DISEASES: text mining and data integration of disease-gene associations. (bioRxiv). <http://dx.doi.org/10.1101/008425>.
- Shao, D.D., Tsherniak, A., Gopal, S., Weir, B.A., Tamayo, P., Stransky, N., Schumacher, S.E., Zack, T.I., Beroukhi, R., Garraway, L.A., et al. (2013). ATARIS: computational quantification of gene suppression phenotypes from multisample RNAi screens. *Genome Res.* 23, 665–678.
- Sheu, J.J., Choi, J.H., Guan, B., Tsai, F.J., Hua, C.H., Lai, M.T., Wang, T.L., and Shih, I.M. (2013). Rsf-1, a chromatin remodeling protein, interacts with cyclin E1 and promotes tumour development. *J. Pathol.* 229, 559–568.
- Shih, I.M., Sheu, J.J., Santillan, A., Nakayama, K., Yen, M.J., Bristow, R.E., Vang, R., Parmigiani, G., Kurman, R.J., Trope, C.G., et al. (2005). Amplification of a chromatin remodeling gene, Rsf-1/HBXAP, in ovarian carcinoma. *Proc. Natl. Acad. Sci. USA* 102, 14004–14009.

- Silva, J.M., Marran, K., Parker, J.S., Silva, J., Golding, M., Schlabach, M.R., Elledge, S.J., Hannon, G.J., and Chang, K. (2008). Profiling essential genes in human mammary cells by multiplex RNAi screening. *Science* 319, 617–620.
- Stephens, P.J., Tarpey, P.S., Davies, H., Van Loo, P., Greenman, C., Wedge, D.C., Nik-Zainal, S., Martin, S., Varela, I., Bignell, G.R., et al.; Oslo Breast Cancer Consortium (OSBREAC) (2012). The landscape of cancer genes and mutational processes in breast cancer. *Nature* 486, 400–404.
- Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* 102, 15545–15550.
- UniProt Consortium (2013). Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nucleic Acids Res.* 41, D43–D47.
- Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., Jr., and Kinzler, K.W. (2013). Cancer genome landscapes. *Science* 339, 1546–1558.
- Walter, V., Nobel, A.B., and Wright, F.A. (2011). DiNAMIC: a method to identify recurrent DNA copy number aberrations in tumors. *Bioinformatics* 27, 678–685.
- Weinstein, I.B., and Joe, A. (2008). Oncogene addiction. *Cancer Res.* 68, 3077–3080, discussion 3080.
- Widmer, C., and Rätsch, G. (2011). Transfer learning in computational biology. *JMLR Workshop Conf Proc.* 27, 207–216.
- Wong, W.C., Kim, D., Carter, H., Diekhans, M., Ryan, M.C., and Karchin, R. (2011). CHASM and SNVBox: toolkit for detecting biologically important single nucleotide mutations in cancer. *Bioinformatics* 27, 2147–2148.
- Yang, S., Zhang, J.J., and Huang, X.Y. (2012). Mouse models for tumor metastasis. *Methods Mol. Biol.* 928, 221–228.
- Yuan, X., Zhang, J., Zhang, S., Yu, G., and Wang, Y. (2012). Comparative analysis of methods for identifying recurrent copy number alterations in cancer. *PLoS ONE* 7, e52516.
- Zack, T.I., Schumacher, S.E., Carter, S.L., Cherniack, A.D., Saksena, G., Tabak, B., Lawrence, M.S., Zhang, C.Z., Wala, J., Mermel, C.H., et al. (2013). Pan-cancer patterns of somatic copy number alteration. *Nat. Genet.* 45, 1134–1140.

TRIP12 and UBR5 Suppress Spreading of Chromatin Ubiquitylation at Damaged Chromosomes

Thorkell Gudjonsson, Matthias Altmeyer, Velibor Savic, Luis Toledo, Christoffel Dinant, Merete Grøfte, Jirina Bartkova, Maria Poulsen, Yasuyoshi Oka, Simon Bekker-Jensen, Niels Mailand, Beate Neumann, Jean-Karim Heriche, Robert Shearer, Darren Saunders, Jiri Bartek, Jiri Lukas,* and Claudia Lukas

*Correspondence: jiri.lukas@cpr.ku.dk

<http://dx.doi.org/10.1016/j.cell.2014.11.027>

(Cell 150, 697–709; August 17, 2012)

Our paper demonstrated that spread of ubiquitylation at chromatin following DNA damage is limited by the proteins TRIP12 and UBR5. In Figure 2C, we presented evidence in support of this by carrying out immunofluorescence on damaged cells treated with siRNAs targeted to TRIP12 and UBR5 mRNAs. We have realized that the upper-left image of Figure 2C contains an inadvertent error. Instead of portraying cells stained for RNF168, the E3 ligase, we portrayed cells stained for RNF168's enzymatic product, conjugated ubiquitin. This image showed the expected phenotype (small foci of modified chromatin) but did not merge with the corresponding fields of cells stained for 53BP1 and DNA (DAPI). This has now been corrected by replacing the upper-left image (green channel) by a field of cells from the same experiment stained directly for the RNF168 protein. Although this error does not affect any of the conclusions or interpretation of these data (RNF168 accumulation and ubiquitylation of its targets are coupled and are cytologically very similar), we apologize for any confusion it might have caused. The corrected figure appears below.

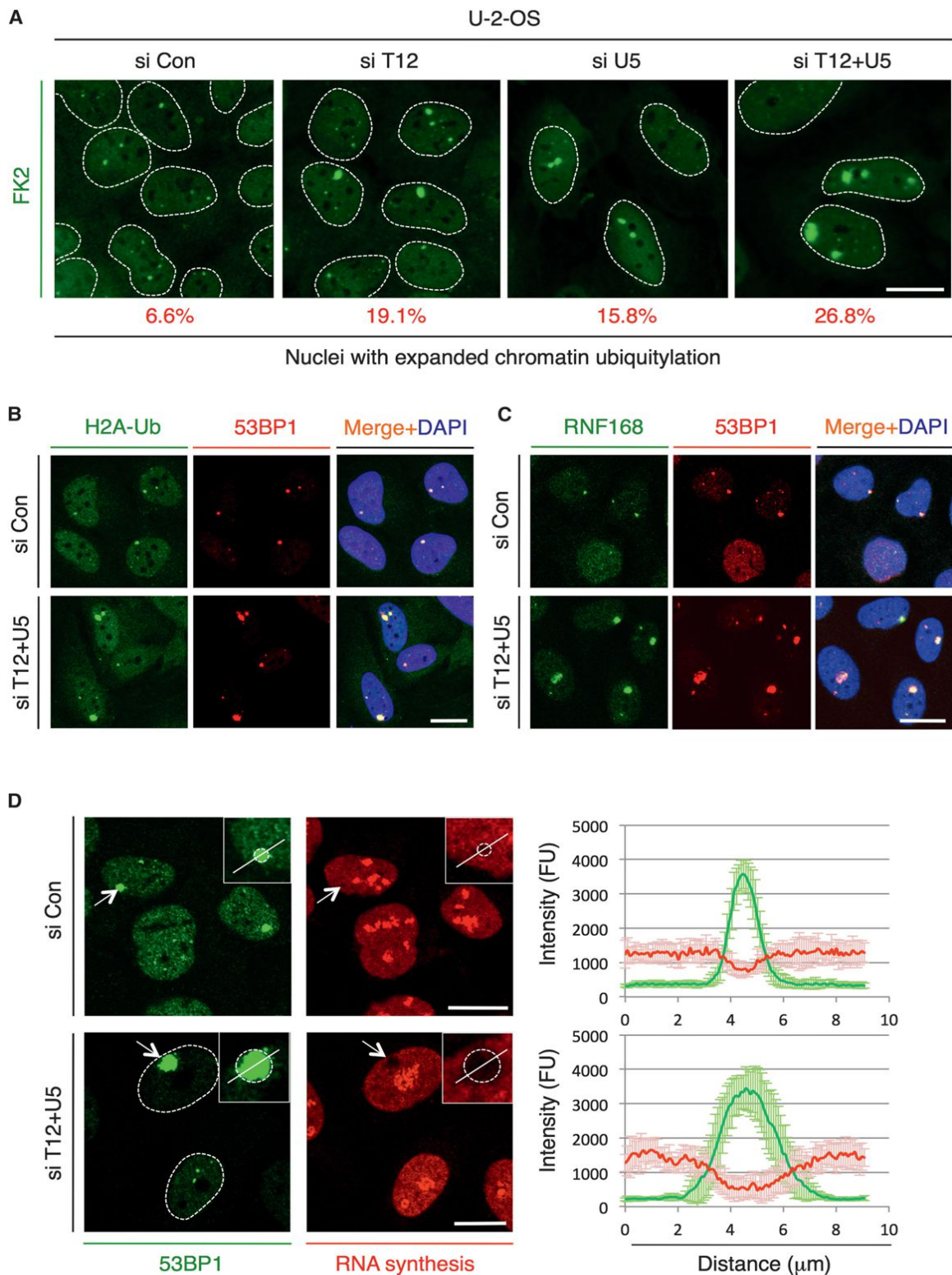


Figure 2. Expansion of the Ubiquitylated Nuclear Domains in TRIP12- and UBR5-Depleted Cells Is Accompanied by Reduced Transcription

Become a Member of ASHG

multi-year discounts now available

Membership benefits include:

- Reduced fees for ASHG Annual Meetings
- Subscription to *The American Journal of Human Genetics*
- Influence on public policy issues
- Access to educational materials
- Free access to invited session recordings from past ASHG Annual Meetings through the membership web portal
- Discount to journal subscriptions (*Nature*, *Nature Reviews Genetics*, and more!)

Regular Members of ASHG get a 20% discount on Regular Membership with the European Society of Human Genetics!



If you have questions about membership, please contact us:

The American Society of Human Genetics
9650 Rockville Pike • Bethesda, MD 20814-3998
1.866.HUMGENE • Membership: scelia@ashg.org



For more information visit:
www.ashg.org/join



Where **Science** Meets the **City**

Abstract Submission Opens:

April 2015

October 6-10 • Baltimore, MD

Abstract Submission Deadline:

June 11, 2015

2015 Program Chair:

Chris Gunter, Emory University

ashg.org/2015meeting

2014 President:

Neil Risch, University of California, San Francisco

The American Society of Human Genetics • 9650 Rockville Pike, Bethesda, MD 20814
1.866.HUMGENE • ashgmeetings@ashg.org



UNIVERSITY OF UTAH
HEALTH SCIENCES

University of Utah
Huntsman Cancer Institute
Department of Oncological Sciences

The **Department of Oncological Sciences** and **Huntsman Cancer Institute** (HCI) invite applications for an **associate professor** or **professor** in the tenure track. We seek an outstanding PhD, MD, or MD/PhD cancer biologist with a well-established laboratory-based research program centered on cancer mechanisms and/or tumorigenesis models, with translational focus or potential. Mechanistic topics might include cancer genetics, cancer cell metabolism, cell cycle control, signal transduction, cell death, gene expression, epigenetics, DNA repair, tumor microenvironment, metastasis, and oncogene/tumor suppressor function. Translational emphasis might include molecular diagnostics, preclinical tumor models, and investigational therapeutics. Departmental strengths include transcriptional regulation, epigenetics, human cancer genetics, stem cell biology, mouse, zebrafish and fly models of cancer, cell signaling pathways, apoptosis, DNA repair, cell motility, and cancer metabolism. HCI is an NCI-designated cancer center with state-of-the-art laboratories and shared resources, including core facilities for imaging, genomics, drug screening and in vivo pre-clinical testing, and population studies. We offer a collegial and interactive research environment and robust graduate programs for training PhD and MD/PhD students.

Review of applications will continue until the position is filled or the search is closed.

For full consideration please apply at the following link:
<http://utah.peopleadmin.com/postings/36883>

Inquiries may be submitted to Recruitment Director at
hci.recruitment@hci.utah.edu

The University of Utah Health Sciences Center is a patient focused center distinguished by collaboration, excellence, leadership, and Respect. The University of Utah HSC values candidates who are committed to fostering and furthering the culture of compassion, collaboration, innovation, accountability, diversity, integrity, quality, and trust that is integral to the mission of the University of Utah Health Sciences Center.

The University of Utah is an Affirmative Action/Equal Opportunity employer and does not discriminate based upon race, national origin, color, religion, sex, age, sexual orientation, gender identity/expression, status as a person with a disability, genetic information, or Protected Veteran status. Individuals from historically underrepresented group, such as minorities, women, qualified persons with disabilities and protected veterans are encouraged to apply. Veterans' preference is extended to qualified applicants, upon request and consistent with University policy and Utah state law. Upon request, reasonable accommodations in the application process will be provided to individuals with disabilities. To inquire about the University's nondiscrimination or affirmative action policies or to request disability accommodation, please contact: Director, Office of Equal Opportunity and Affirmative Action, 201 S. Presidents Circle, Rm 135 (801) 581-8365.

The University of Utah values candidates who have experience working in settings with students from diverse backgrounds, and possess a demonstrated commitment to improving access to higher education for historically underrepresented students.



JOB POSTINGS FROM

40080 SPECIALTIES

SOCIETY AND ASSOCIATION JOURNALS



Announcements/Positions Available careers.cell.com



THE
AMERICAN
SOCIETY
OF HUMAN
GENETICS

Join or Renew Your ASHG Membership Today to receive:

- Reduced Annual Meeting Registration Rates
- Subscription to *AJHG*
- No fee for articles published in *AJHG*

ashg.org/join



THE OHIO STATE UNIVERSITY

Faculty Position in Cancer Biology

The Solid Tumor Biology Program at the Ohio State University Comprehensive Cancer Center and the Department of Molecular Genetics in the College of Arts and Sciences invites applications for tenure-track faculty positions at the levels of Assistant, Associate and Full Professor. Outstanding individuals using and/or developing animal models of endocrine cancer are particularly encouraged to apply. At present, research in the program and department spans a wide range of topics in cancer biology including genetics, metabolism, signaling, cell cycle, cell differentiation, DNA repair, transcription, DNA replication, checkpoint control, cachexia, aging, and tumor microenvironment. Additional information about the OSUCCC and the Department of Molecular Genetics is available at <http://cancer.osu.edu> and <http://molgen.osu.edu>.

Applicants should submit a cover letter, curriculum vitae, and a statement of research and teaching interests, and three letters of reference. Inquiries may be directed to Jessie Siegman (Siegman.1@osu.edu). Review of applications will begin on December 15, 2014, and we encourage submission before that date. However, applications will continue to be accepted until February 15, 2015. Please apply online through Academic Jobs Online at: <http://academicjobsosonline.org>.

The Ohio State University is an equal opportunity employer. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, sexual orientation or identity, national origin, disability status, or protected veteran status. Ohio State is an NSF Advance Institution.

A PARTNERSHIP BETWEEN

James Cancer Hospital and Solove Research Institute
OSU Comprehensive Cancer Center, Solid Tumor Biology Program
OSU College of Arts and Sciences, Department of Molecular Genetics



Be their next big discovery.

careers.cell.com

Where top life sciences talent is searching for their next opportunity

Cell Career Network

Discover the best jobs in life science

To showcase your job openings on Cell Career Network, contact: advertising@cell.com

SnapShot: Hormones of the Gastrointestinal Tract

Cell

Katie C. Coate,^{1,2} Steven A. Klierer,^{1,3} and David J. Mangelsdorf^{1,2}
¹Department of Pharmacology, ²Howard Hughes Medical Institute, and
³Molecular Biology, University of Texas Southwestern Medical Center, Dallas, TX 75235, USA

STOMACH					SMALL and/or LARGE INTESTINE				
Hormone	Site(s) of Production	Stimulus for Secretion	Receptor / Site of Action	Major Function(s)					
Gastrin	Primarily in G cells of gastric antrum; variable extent in the duodenum; much lesser extent in colon and pancreas	Food ingestion (primarily protein)	CCK2 receptor (a GPCR) on fundic enterochromafin-like cells and gastric parietal cells	Stimulates gastric acid secretion and epithelial cell proliferation; participates in iron homeostasis					
Ghrelin	Primarily in X/A-like (in rodents) or P/D1 (in humans) endocrine cells of the oxyntic mucosa	Fasting, before a meal	Centrally through growth hormone secretagogue receptor-1a (a GPCR) and peripherally through its receptor on vagal afferents innervating the stomach	Orexigenic effect on appetite/feeding; stimulates gastric emptying, acid secretion, and migrating motor complexes; protects against gastric stress; increases release of growth hormone					
Leptin	Chief cells and endocrine P cells in the gastric fundic region	Food ingestion, vagal nerve stimulation, CCK, and secretin	Leptin receptor Ob-R (in the gp130 family of cytokine receptors) on gastric vagal afferents and on the apical side of enterocytes along the small and large intestine	Anorexigenic effect on appetite/feeding; stimulates CCK and GLP-1 secretion; modulates intestinal absorption of nutrients					
Somatostatin	Enteroendocrine D cells in the antral and fundic mucosa of the stomach and along the intestinal mucosa	Intraluminal nutrients and acid, adrenergic stimulation, CCK, and gastrin	Somatostatin receptor (rhodopsin-like GPCR, mostly the SST ₂ subtype) along the GI tract and in the pancreas	Inhibits gastric acid secretion and endocrine and exocrine pancreatic secretion					
Secretin	Throughout the small intestine but primarily in enteroendocrine S cells of duodenal mucosa	Acidic chyme from stomach, digested fat and protein	Secretin receptor (family B GPCR) located on basolateral membrane of ductal and centroacinar cells of pancreas, on epithelial cells of large intrahepatic bile duct units, and in the kidney	Stimulates alkaline secretion from the pancreas (bicarbonate, water, and electrolytes) and biliary ductular systems; inhibits gastric motility and acid secretion; participates in body fluid homeostasis / osmoregulation					
GIP	Enteroendocrine K cells of the proximal small intestine (mainly duodenum)	Food ingestion (primarily carbohydrates and fat)	GIP receptor (a GPCR) on the endocrine pancreas	Acts as an incretin hormone that potentiates glucose-stimulated insulin secretion					
CCK	Enteroendocrine I cells of duodenum and jejunum	Food ingestion (primarily protein and fat)	CCK1, and perhaps CCK2 receptors (GPCRs), on vagal afferents, the stomach and upper small intestine, the pancreas and gallbladder, and in the CNS	Reduces feeding/meal size; inhibits gastric emptying and acid secretion; stimulates gall bladder contraction and pancreatic digestive enzyme secretion					
Motilin	Enteroendocrine M cells of the proximal small intestine	Interdigestive fasting period	Motilin receptor (a GPCR) on nerves and muscle of the GI tract	Induces phase III contraction of the migrating motor complex (GI motor activity)					
PYY	Enteroendocrine L cells of the jejunum, ileum, and colon	Food ingestion (primarily fat)	At least five distinct Y receptor subtypes of the GPCR family (Y1, Y2, Y4, Y5, Y6) along the GI tract, in the pancreas, and in the CNS	Reduces food intake; inhibits gastric emptying and secretion; suppresses intestinal motility and electrolyte secretion; inhibits pancreatic secretion					
GLP-1	Enteroendocrine L cells of jejunum, ileum, and colon	Food ingestion (particularly carbohydrates and fat); bile acids acting on TGR5 (a GPCR)	GLP-1 receptor (a GPCR) along the GI tract, in the endocrine pancreas, on vagal afferents, and in the CNS	Incretin effect (potentiates glucose-stimulated insulin secretion); reduces food intake; inhibits gastric emptying and GI secretion; inhibits glucagon secretion					
GLP-2	Enteroendocrine L cells of jejunum, ileum, and colon	Co-secreted with GLP-1 in response to nutrient ingestion	GLP-2 receptor (a GPCR) in the GI tract and the enteric and CNS	Stimulates cell growth in the gut mucosa and protects against apoptosis; inhibits gastric emptying and acid secretion; enhances intestinal nutrient absorption and blood flow					
OXM	Enteroendocrine L cells of jejunum, ileum, and colon	Food ingestion (particularly fat)	GLP-1 and glucagon receptors (GPCRs)	Reduces food intake; reduces gastric acid and exocrine pancreatic secretion; potentiates glucose-stimulated insulin secretion					
FGF19 (human) FGF15 (rodent)	Ileal enterocytes	Bile acids acting on FXR (a nuclear receptor)	Heteromeric receptor comprised of EGF receptor 4 (a tyrosine kinase receptor) and b-klotho (a single transmembrane co-receptor) in the liver	Suppresses bile acid synthesis; stimulates hepatic protein synthesis and glycogenesis; suppresses gluconeogenesis					

Supported by:



Even more ways to access Cell Press journals



App Features:

- Available to subscribers at no additional charge
- Free 30-day trial to all 30 Cell Press journals
- Interactive reading experience
- Browse abstracts and full-text articles
- Download now, read later
- Make article notes
- Share with your friends and colleagues
- Now on the iPhone, iPad, Android tablet and Android phone



Download the app at
cell.com/mobile

CellPress



Have your cake and eat it too **R&D Systems Luminex[®] Assays**

Get the **most plexible analytes** and enjoy
the **accuracy and precision you expect**
from R&D Systems products.

Find out how to win a cake for your lab.
rndsystems.com/LabCake

R&D systems[™]
a biotechnne brand